MULTISCALE MODELING AND MACHINE LEARNING STUDIES OF THE DIFFUSION OF SILICON AND INTRINSIC DEFECTS IN III-V SEMICONDUCTORS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Mardochee Reveil

August 2018

MULTISCALE MODELING AND MACHINE LEARNING STUDIES OF THE DIFFUSION OF SILICON AND INTRINSIC DEFECTS IN III-V SEMICONDUCTORS

Mardochee Reveil, Ph.D.

Cornell University 2018

Integrating III-V semiconductors into next-generation silicon-based processing is a promising alternative being considered as a route to faster and more energy-efficient electronic devices. These III-V materials will be doped, typically with Si as a dopant. However, dopant activation remains an issue, compounded by the fact that there is still insufficient knowledge of the ease and preferred mechanistic pathways by which dopants, like Si, become activated within the III-V matrix. Using Density Functional Theory (DFT) and Nudge Elastic Band (NEB) calculations, we have determined many of these critically important properties, namely, the energy barriers associated with the diffusion of both intrinsic point defects and silicon impurities in prototypical III-V materials, here Zinc Blende GaAs, InAs and the CuAuI-ordered ternary $In_{0.5}Ga_{0.5}As$.

Refuting assumptions in the current literature that the enhanced diffusion can be attributed primarily to an increase in vacancies, vacancy-assisted diffusion of isolated Si atoms was found to be an unfavorable mechanism for this group of semiconductor alloys. Our results show that new and highly mobile species that are created at high dopant concentration are instead responsible for the enhanced diffusion observed. Those new species include Si complexes such as Si-Si pairs and Si split interstitials which can move more easily within the crystal lattice. We use these DFT results to inform the development of a

continuum model that addresses limitations in current models and shows great agreement with experimental results. We also develop a new method whereby machine learning in lieu of DFT, is used to predict forces during NEB simulations. This new method allows to compute transition pathways at a fraction of the cost while maintaining reasonable accuracy compared to the traditional DFT approach.

BIOGRAPHICAL SKETCH

Mardochee Reveil was born and raised in Haiti where he completed his Bachelors degree in Chemistry before moving to the United States in 2011. He earned a Masters Degree in Chemical Engineering at Syracuse University in 2014 and interned twice with Intel Corporation in Oregon. During his first internship with Intel, he developed a special interest in scientific research and decided to shift his focus to pursue a PhD. He later joined Cornell University where he started his doctoral research under Prof. Paulette Clancy in the Chemical Engineering department with a focus on studying diffusion in semiconductors using advanced molecular simulations and machine learning techniques. In 2016, he interned with Corning Incorporated and joined the company as a Research Scientist in January 2017 while concurrently completing his PhD degree. At Corning, he works on several fundamental research projects related to glass composition design and value-enhancing glass processing techniques.

This thesis is dedicated to my Mother, Antoinette Dormévil who through her life and teachings, has instilled in me the values of hard work, integrity, passion and generosity.

ACKNOWLEDGEMENTS

Completing this PhD program would not have been possible without the help and support of exceptional individuals who have contributed in one way or another towards my success.

First, I would like to thank my advisor, Dr. Paulette Clancy, for her incredible mentoring and support. I truly consider myself lucky to have had the opportunity to work in her group over the past four years. I'm especially grateful to have enjoyed, under her leadership, a significant amount of freedom not only in the types of problems that I studied, but also in how I approached them. Her open-mindedness and expert support have been instrumental in helping me grow as a researcher. During this time, I've also gotten to know her as a wonderful person with a big heart. She truly cares about her students and displays unwavering support towards both their personal and professional growth.

Next, I would like to thank my committee members, Dr. Mike Thompson, Dr. Fernando Escobedo and Dr. Perrine Pepiot. Dr. Thompson's expertise in fundamental theory and processing of semiconductors was extremely useful in helping me complete my projects. I would like to also thank my support network at Corning Incorporated, specially Lisa Hepburn, Dr. John Mauro and Dr. Adama Tandia who played an important role in helping me manage my professional responsibilities while completing my PhD research in a timely fashion.

I also want to thank members of the Clancy group who maintained a friendly, supportive and collegial atmosphere within the group allowing every member to feel welcome and at home. I'd like to especially thank Dr. Victoria Sorg who helped me early on in my PhD studies to get started with my projects. I'd like to thank Dr. Jonathan Saathoff, Dr. James Stevenson, Jingyang Wang, Yaset Acevedo, Henry Herbol and Nikita Sengar for countless useful dis-

cussions, exchange of ideas and tremendous support throughout my studies. I'd like to also thank my mentees, Jason Liu, Hsien-Lien Huang and Huang-Ta Chen whom I had the pleasure to work with and who in turn mentored me on how to better mentor.

I've also been fortunate to have wonderful colleagues and friends who have been extremely supportive throughout my PhD studies. Special thanks to Dr. Bendy Estime, Obste Therasme, Wolf Peter Jean-Philippe, Dr. Kofi Appiah-Okyere, Jane Appiah-Okyere and Patricia Birago to name a few. Their friendship and moral support are priceless and helped me tremendously.

I want to thank my family, especially my Mom and Dad who made this accomplishment possible through their countless sacrifices over the years to raise and provide me with a good education. I'd like to also thank my sisters, Naomi-Esther Reveil and Rose-Merline Reveil for their tremendous support. Last but not least, I want to thank my girlfriend, Lory Henderson, who has been by my side over the past few years and whose intelligence, beauty and grace never cease to captivate my heart and inspire me to do my best.

CONTENTS

Biographical Sketch . . . iii
Dedication . . . iv
Acknowledgements . . . v
Contents . . . vii
List of Tables . . . x
List of Figures . . . xii

1 Introduction **1**
1.1 Overview of semiconductors . . . 1
1.2 III-V semiconductors . . . 3
1.3 Dopants and defects in semiconductors . . . 4
1.4 Silicon as a dopant in III-Vs . . . 7
1.5 Research objectives . . . 10
1.6 Thesis outline . . . 12

2 Methods **13**
2.1 Overview . . . 13
2.2 Density Functional Theory . . . 13
2.3 Nudged Elastic Band . . . 17
2.4 Machine Learning . . . 20
2.5 Artificial Neural Networks . . . 23

3 Intrinsic and silicon diffusion in GaAs **27**
3.1 Introduction . . . 27
3.2 Diffusion results from Literature . . . 30
3.2.1 Silicon . . . 30
3.2.2 Other group IV dopants: Tin and Germanium . . . 31
3.2.3 Overview of existing analytical diffusion models . . . 33
3.3 Methods . . . 34
3.4 Results and discussion . . . 37
3.4.1 Vacancy-assisted diffusion . . . 37
3.4.2 Interstitial impurities . . . 43
3.5 Conclusions . . . 46

4 Intrinsic and silicon diffusion in InAs **47**
4.1 Introduction . . . 47
4.2 Methods . . . 48
4.3 Results and discussion . . . 52
4.3.1 Vacancy-assisted diffusion . . . 52
4.3.2 Interstitial diffusion . . . 57
4.3.3 Defect formation energies . . . 59
4.4 Conclusions . . . 69

5 Intrinsic and silicon diffusion in InGaAs **72**
5.1 Introduction . . . 72
5.2 Methods . . . 73
5.2.1 Crystal structure and ordering . . . 73
5.2.2 Mechanisms studied . . . 75
5.2.3 Simulation details . . . 75
5.3 Results and discussion . . . 78
5.3.1 Vacancy-assisted diffusion . . . 78
5.3.2 Interstitial diffusion . . . 85
5.4 Conclusions . . . 94

6 A continuum model for amphoteric dopant diffusion in III-Vs **95**
6.1 Introduction . . . 95
6.1.1 Limitations of the Greiner and Gibbons model . . . 95
6.1.2 Limitations of the Yu *et al.* model . . . 96
6.2 Continuum model derivation . . . 97
6.3 Results and Discussion . . . 104
6.3.1 Concentration of Si species: An example in GaAs . . . 104
6.3.2 Effective Si diffusivity in GaAs . . . 105
6.3.3 Comparison with previous models . . . 112
6.3.4 Diffusion profile in InGaAs . . . 114
6.3.5 Limitations of the Model . . . 114
6.4 Conclusions . . . 117

7 Molecular fingerprints for machine learning applications **119**
7.1 Introduction . . . 119
7.2 Overview ot fingerprinting strategies . . . 120
7.3 Classification of fingerprints . . . 121
7.4 Review of key fingerprints . . . 126
7.4.1 Symmetry Function Fingerprints . . . 126
7.4.2 Bispectrum Fingerprints . . . 129
7.4.3 Zernike Fingerprints . . . 130
7.4.4 SOAP Fingerprints . . . 131
7.5 Implementation of an open-source package for fingerprint computation . . . 132

8 Machine-learning determination of diffusion pathways **138**
8.1 Introduction . . . 138
8.2 Method description . . . 140
8.3 Simulation details . . . 144
8.4 Results and discussion . . . 148
8.4.1 Effect of dataset size and network architecture . . . 148
8.4.2 Machine learning models for force prediction . . . 149
8.4.3 Predictions of diffusion pathways . . . 153

8.5 Limitations and future work 155
8.6 Conclusions 158

9 Limitations and future work **159**
9.1 Experimental validation 159
9.2 Modeling of random InGaAs crystals 160
9.3 Beyond the In-Ga-As family of III-V materials 161

10 Conclusions **162**

Bibliography **164**

LIST OF TABLES

3.1 List of vacancy-assisted transitions studied in this work. 36

4.1 Summary of energy barrier results from DFT NEB calculations for vacancy-assisted transitions in InAs. N/A refers to transitions that do not take place because it would involve a monotonic increase in energy of the system (with no barrier). Conversely, "No barrier" refers to a transition that goes downhill in the energy landscape. 58

4.2 Summary of energy barrier results from DFT NEB calculations for interstitial transitions in InAs showing the generally lower barriers for interstitial jumps. 68

5.1 List of vacancy-assisted transitions studied in this work. 76

5.2 Summary of energy barrier results from DFT NEB calculations for first-nearest neighbor vacancy-assisted transitions in CuAuI-ordered InGaAs. "N/A" indicates a transition that does not take place as it would lead to a high energy configuration without involving an energy barrier. The reverse case is labeled "No barrier" and refers to a transition that takes place spontaneously from a high energy configuration to a low energy configuration with no energy barrier. 81

5.3 Summary of energy barriers for second-nearest neighbor vacancy-assisted transitions in CuAuI-ordered InGaAs. The superscript denotes the common nearest neighbor between the two As lattice sites in the case of As and Si jumps. 82

5.4 Summary of energy barriers for interstitial transitions in CuAuI InGaAs. Two different paths for S_1 to S_1 Si transitions were explored. 91

6.1 List of parameter values obtained from fitting our diffusion model to experimental data for Si diffusion in GaAs assuming constant α values. 108

6.2 List of parameter values obtained from fitting our diffusion model to experimental data for Si diffusion in GaAs assuming concentration-dependent α values. Again we assume $K \to 0$. . . 110

6.3 List of parameter values used to fit our diffusion model to experimental data for Si diffusion in InGaAs at 750 C. 115

7.1 Overview of the main fingerprinting strategies for machine learning applications proposed over the past ten years using the classification scheme developed in this paper. 123

8.1 Parameter values used to compute the fingerprints for machine learning force prediciton in InGaAs. For more details on those fingerprints and the meaning of the parameters see Section 7.4. . 145
8.2 Fingerprint size for each of the eight types of machine learning models trained in this work. To speed up the process, a parallel processing approach is adopted whereby a specific model is trained for each atom type (In, Ga or As) and each force direction (x, y or z). This lead to a total of 72 different models. 147
8.3 List of the transitions selected for comparison between DFT and ML NEB diffusion results. 156

LIST OF FIGURES

1.1 Example of a subsitutional dopant (shown in tan) and a vacancy defect in a 64-atom InGaAs crystal in a Zinc Blende structure and CuAu-I ordering. As, In and Ga are respectively shown in red, yellow and green. The black dashed arrow denotes a possible transition whereby the substitutional defect would move to the nearby vacancy site. 6

2.1 Example energy profile along an atomistic jump or chemical reaction. The energy barriers give access to the forward and backward rates. 19

2.2 (a) Schematic representation of a the Perceptron. This basic implementation of a Neural Network has only one neuron and uses a Heaviside step function to transform a linear combination of the features of the input vector into a label(0 or 1) that can be used for binary classification. (b) Schematic representation of an artificial Neural Network. This can be seen as an extension of the Perceptron whereby multiple neurons are combined into layers allowing to learn more complicated functions. 25

3.1 Number of articles published between 1975 and 2018 with InGaAs, GaAs or InAs in their title or list of keywords (data from SCOPUS - www.scopus.com). 28

3.2 Example of a 64-atom GaAs crystal in the Zinc Blende structure with As shown in red and Ga in green. 29

3.3 Si diffusion by first nearest neighbor vacancy-assisted jumps. . . 38

3.4 Energy barriers for Si diffusion by second nearest-neighbor vacancy-assisted jumps. 39

3.5 $Si_{Ga} - Si_{As}$ pair undergoing a vacancy-assisted transition on the Ga sublattice in a 64-atom GaAs supercell. Ga, In and Si are respectively shown in green, red and blue. 41

3.6 Energy barriers for Si-Si pairs diffusing by second nearest-neighbor vacancy-assisted jumps. 42

3.7 Optimized geometry for a As atom (in yellow) in a split interstitial configuration in GaAs. Ga atoms are in green and As atoms in red. 44

3.8 Optimized geometry for a Si atom (in blue) in a split interstitial configuration in GaAs. Ga atoms are in green and As atoms in red. 45

4.1 A representative 64-atom InAs supercell in the cubic Zincblende structure. Unless otherwise noted, indium atoms are shown in yellow and arsenic in red. 50

4.2 DFT-NEB results showing an energy barrier of 1.6 eV for In to jump to a nearby In vacancy in the InAs lattice. For improved clarity, the moving In atom is shown in green in the insets. 54
4.3 DFT-NEB results for vacancy-assisted As jumps in InAs. Arsenic can undergo transitions that takes it to both In and As vacancies alike. The insets show the initial, saddle point and final configurations of a first nearest-neighbor As jump (as represented by the green curve). The moving As atom is shown in blue for added clarity. 55
4.4 DFT-NEB results for vacancy-assisted Si jumps in InAs. The insets show the initial, saddle point and final configurations of Si (colored in blue for added clarity) moving between two In sites (green curve). 56
4.5 An indium atom (shown in green) in tetrahedral interstitial configuration in InAs. 60
4.6 An arsenic split-interstitial configuration in InAs. 61
4.7 A silicon atom colored in blue in hexagonal interstitial position in InAs. 62
4.8 DFT-NEB results for In interstitial jumps in InAs showing a relatively low 0.8 eV barrier. The moving In atom is colored in green.The saddle point corresponds to an hexagonal configuration. 63
4.9 DFT-NEB results for As interstitial jumps in InAs, showing the ease of As diffusion in InAs (< 0.5 eV barrier) *via* an interstitial mechanism. The As atoms involved in the transition are shown in green for more clarity. 64
4.10 DFT-NEB results for Si interstitial jumps in InAs, showing the lowest energy barriers among the set of pathways we investigated, a < 0.25 eV barrier for hexagonal transport (blue curve) and showing that the hexagonal interstitial site is preferred to the tetrahedral one (green curve) by about 0.2 eV.The hexagonal to tetrahedral transition is illustrated by insets showing the initial, saddle point and final configurations with the moving Si atom colored in blue. 65
4.11 DFT-NEB results for In interstitial jumps in InAs showing a relatively low 0.8 eV barrier. The moving In atom is colored in green.The saddle point corresponds to an hexagonal configuration. 66
4.12 DFT-NEB results for As interstitial jumps in InAs, showing the ease of As diffusion in InAs (< 0.5 eV barrier) *via* an interstitial mechanism. The As atoms involved in the transition are shown in green for more clarity. 67
4.13 Formation energy of point defects in InAs with respect to growth conditions. 70

5.1 (a) Schematic of a 64-Atom CuAuI-ordered $In_{0.5}Ga_{0.5}As$ supercell, showing alternating layers of atoms. The CuAuI ordering of InGaAs has been observed for specific growth conditions, *e.g.*, MBE growth. (b) Color key used throughout this text. From left to right: As, Si, Ga and In (Images not on scale) 74
5.2 Energy profiles for As moving from a Ga site to an As vacancy with (green dotted line) and without (blue solid line) a Si atom as second nearest neighbor to a Ga site. The presence of Si renders the As_{Ga} defect unstable. The curve with silicon corresponds to a geometry optimization of the initial configuration which is shown to relax to the final configuration with no barrier. The curve without silicon corresponds to an actual NEB calculation as this transition involve an energy barrier. 84
5.3 The three possible interstitial sites in CuAuI-ordered $In0.5Ga0.5As$. Left: T_{1a} with four As as first nearest neighbors (NN), four Ga and two In as second NN. Center: T_{1b} with four As as first NN, four In and two Ga as second NN. Right: T_2 with two In and two Ga as first NN, six As as second NN. 87
5.4 Examples of As split interstitial configurations. Left: S_1 configuration with two As atoms sharing an As lattice site. Center: S_2 configuration with two As atoms sharing a lattice site with a slight offset in the z direction. Right: S_3 configuration with an As sitting in an hexagonal position leading to slightly displaced In atom. 88
5.5 Examples of Si split interstitial configurations. Left: S1 configuration with a Si and an As atom sharing an As lattice site. Right: S3 configuration with a Si sitting in an hexagonal position leading to slightly displaced In atoms. 89
5.6 Formation energy of Si split interstitials ($S1$ and $S3$) and tetrahedral interstitials ($T1a$, $T1b$, $T2$) in $In_{0.5}Ga_{0.5}As$. 93

6.1 Concentration of the different species of Si atoms of interest in our continuum model of diffusion as a function of total Si concentration and the parameters K and K_p. The influence of K_p on the relative distribution of Si atoms into Si_{Ga}, Si_{As} and Si pairs is significant. 106
6.2 Our effective diffusivity model fitted to experimental data for Si diffusion in GaAs for constant α values 109
6.3 Our effective diffusivity model fitted to experimental data for Si diffusion in GaAs with concentration-dependent α values. . . . 111
6.4 Effective diffusivity predicted by the Greiner model as a function of total Si concentration. This model is unable to capture the sigmoid-like shape of experimentally observed effective diffusivity as a function of dopant concentration. 113

6.5 Fitting of experimental Si diffusion profile in InGaAs at 750 C with our model. Data from Aldridge Jr. *et al.* [1]. 116

7.1 Two distinct strategies exist to generate fingerpints: global and atom-centered. The former is generally suitable for macroscopic property predictions whereas the latter can be used for local property predictions and are particularly relevant to force field development. 124
7.2 Comparison between the current way of conducting machine learning-based molecular studies (given in dashed lines and highlighted in red) and the proposed method using SEING (highlighted in blue). By providing a unified framework and package for fingerprint generation in an "off-the-shelf" fashion, SEING will make such machine learning studies more straightforward and effective. 134
7.3 Flowchart showing the general procedure used by SEING to compute fingerprints. 136

8.1 Flowchart showing the general procedure used to build a Machine Learning model to predict atomic forces on each atom. . . 141
8.2 Mean Absolute Error (MAE) (a) and R^2 scores (b) of a Neural Network model for force prediction on As atoms as a function of dataset size. The scores are calculated based on a held-out test set of 20,000 data points. The results show that 10,000 points or more are needed to train the Neural Network for proper generalization. 150
8.3 Mean Absolute Error (a), Root Mean Square Error (b) and R^2 scores (c) as a function of number of neurons for force prediction on a retained set of 20,000 data points. Each line represents a network with a different number of layers. 151
8.4 Performance of Neural Network machine learning models for force prediction in InGaAs crystals with Behler-Parinelo fingerprints computed with SEING, our custom-made software package for fingerprint computing. 152
8.5 Flowchart describing the new Machine Learning NEB framework used in this work. This assumes that a suitable Machine Learning force prediction model has already been properly trained.154
8.6 Comparison between the diffusion paths obtained through DFT calculations and our ML method expressed in terms of distance between the two paths (or deviations from the DFT path). 157

CHAPTER 1

INTRODUCTION

1.1 Overview of semiconductors

Semiconductors are a diverse class of materials which, due to their technological importance, play a key role in current human civilization. Their composition space span from simple monoatomic crystals such as silicon [2] and germanium [3] to compounds and complex alloys such as InGaAsN [4], and organic-inorganic structures such as Perovskites [5, 6]. Their application areas are also very diverse and critical to many industries. One prime example of their technological importance is in the microelectronics industry which has upheld a miniaturization trend of CMOS transistors, the semiconductor-based logic devices at the heart of modern computing technologies, for about half a century [7]. This trend called Moores's law [8] in honor of Gordon Moore who first reported it in 1965, is largely responsible for the exponential increase in computing power available to us today in continuously smaller devices. This semiconductor based revolution has in turn revolutionized the way people live, manufacture goods, create, store and exchange information, etc. [9]. Semiconductors are also the core materials used for photo-electrical energy conversion and are therefore used in solar cells [10, 11, 12], lasers[13, 14, 15], Light Emitting Diodes (LEDs) [16]. Other application areas include radars, photodetectors, amplifiers, sensors, etc. [17].

From a modern perspective, semiconductors are classified as such based on the value of their "bandgap" (prohibited region in their electronic energy band diagram) [18, 17]. Electrical conductors such as Cu have small bandgaps al-

lowing electrons to easily move from valence to conduction bands and act as electrical carriers. On the other hand, insulators have large bandgaps that make it difficult for electrons to become charge carriers. Semiconductors, however, have bandgaps that fall in between that of insulators and conductors. As such, they lend themselves to various bandgap and electronic engineering strategies with significant scientific and technological applications.

The probability distribution of electron occupation in a semiconductor as a function of energy E and temperature T follows the Fermi-Dirac statistics given by Eq. 1.1 where E_F is the Fermi energy and k_B is Boltzmann's constant. The Fermi energy is defined as the highest occupied energy level at 0 K.

$$f(E) = \frac{1}{e^{(E-E_F)/k_BT} + 1} \tag{1.1}$$

If we take the valence band maximum as $E = 0$ and make a few additional assumptions such as harmonic parabolic form of the conduction band and high enough energy so that the Fermi distribution can be approximated by $f(E) \approx e^{(E_F-E)/k_BT}$, it can be shown that the electron concentration n as a function of temperature T will follow Eq. 1.2 given below:

$$n = N_c e^{(E_F-E_g)/k_BT} \tag{1.2}$$

where

$$N_C = 2\left(\frac{m_e k_B T}{2\pi\hbar^2}\right)^3 /2 \tag{1.3}$$

is the effective density of states of the conduction band. m_e is the electron effective mass, E_g the bandgap and $\hbar$, Plank's constant. E_F in Eq. 1.2 is the Fermi level and corresponds to the energy level with a 50% probability of occupation. Such quantity has no particular physical meaning but has nonetheless become a key metric to describe the characteristics of a semiconductor and as such, will be used as needed in this thesis. For more in-depth background on semiconductors, the reader is referred to McCluskey and Hallier [19] among other excellent books written on the subject.

1.2 III-V semiconductors

III-V semiconductors are a subgroup of compound semiconductors made of elements from groups III and V in the periodic table [20]. Other compound semiconductors are also possible with elements from groups II and VI, and groups I and VII (respectively called II-VI and I-VII semiconductors) [21, 22, 23, 24]. Those compound semiconductors share the fact that they all have exactly 8 electrons in their outer orbitals. [25] Examples of III-V semiconductors include binaries such as GaAs, InAs, InP, GaN, etc. ternaries such as $In_xGa_{1-x}As$ where $0 < x < 1$, quaternaries such as InGaAsP, etc. [26, 27]. It is apparent here that even the subgroup of semiconductors made of elements from groups III and V span a large composition space.

III-V compounds usually have a direct bandgap. This means electrons can go from valence to conduction band and vice-versa without involving phonons for energy conservation [28, 20]. This usually translates into better efficiency for opto-electronic applications [29]. In addition, their effective electron masses are

usually lower and their electron mobilities higher [28] compared to group IV semiconductors. This makes them highly appropriate for high-efficiency and high-speed amplifiers and transistors applications [28]. Moreover, compared to other compound semiconductors, III-Vs are less polar and therefore easier to be p-doped. They also benefit from less undesirable defect densities prejudicial to electronic and optical applications [30].

Because of such unique properties, integration of III-V materials into transistor technologies currently dominated by silicon, has been identified as an important research area in the International Technology Roadmap for Semiconductors (ITRS) [31]. Among other benefits, such integration would allow to take advantage of the outstanding electron mobilities of those materials. For example, with a mobility at least six times greater than silicon and a direct band gap, $In_{0.5}Ga_{0.5}As$ (InGaAs), a pseudo-binary alloy of InAs and GaAs, is one of the most promising III-V materials for integration into high-performance and energy-efficient future transistors [28]. This thesis is focused on developing fundamental understanding for the family of In, Ga and As-based semiconductors which include the binaries InAs, GaAs and the ternaries $In_xGa_{1-x}As$ where $0 < x < 1$ and more specifically $In_{0.5}Ga_{0.5}As$, which will be called InGaAs for the remainder of this text.

1.3 Dopants and defects in semiconductors

Foreign atoms are usually introduced in semiconductors to modify their electronic properties. Those impurity atoms are called ”dopants” and the process of introducing them, doping [23]. Inside the host crystal, dopants can occupy

lattice sites to become ”substitutional defects” or remain in interstitial positions. In addition to those extrinsic defects caused by the presence of impurities (introduced intentionally or not), other defects, called ”intrinsic defects” can be present in the host crystal [32]. For example, ”vacancies” are created when a typically occupied site is empty. An example vacancy and substitutional defect are shown in Fig. 1.1 Host atoms can also be found in interstitial sites thus forming ”interstitial defects” or occupy the wrong lattice site in which case they are called ”antisites”.

In practice those point-defects provide a ”knob” to manipulate the Fermi level and consequently the concentration of charge carriers which in turn allows to engineer transistors and various other micro/nano-scale semiconductor devices [18, 17]. The goal is usually to create semiconductors which are p-type or n-type based on whether the majority charge carriers are electrons (n-type) or holes (p-type) [18, 17]. For n-type doping, the foreign atoms occupy sites that are normally occupied by atoms with less valence electrons in which case they provide extra electrons for conduction. Conversely, p-type doping involves sites normally occupied by atoms with more valence electrons leading to the creation of holes (or absence of electrons). For example, Boron (group III) would be a p-type dopant in silicon (group IV) whereas Phosphorus (group V) would be an n-type dopant. In compound semiconductors such as the III-Vs, a group IV element such as silicon can occupy both group III sites (which would lead to n-type doping) or group V sites (which would lead to p-type doping) [33]. Silicon and other group IV elements are therefore amphoteric dopants in III-Vs whereas group VI or above atoms are n-type dopants and group II and below are p-type dopants [33].

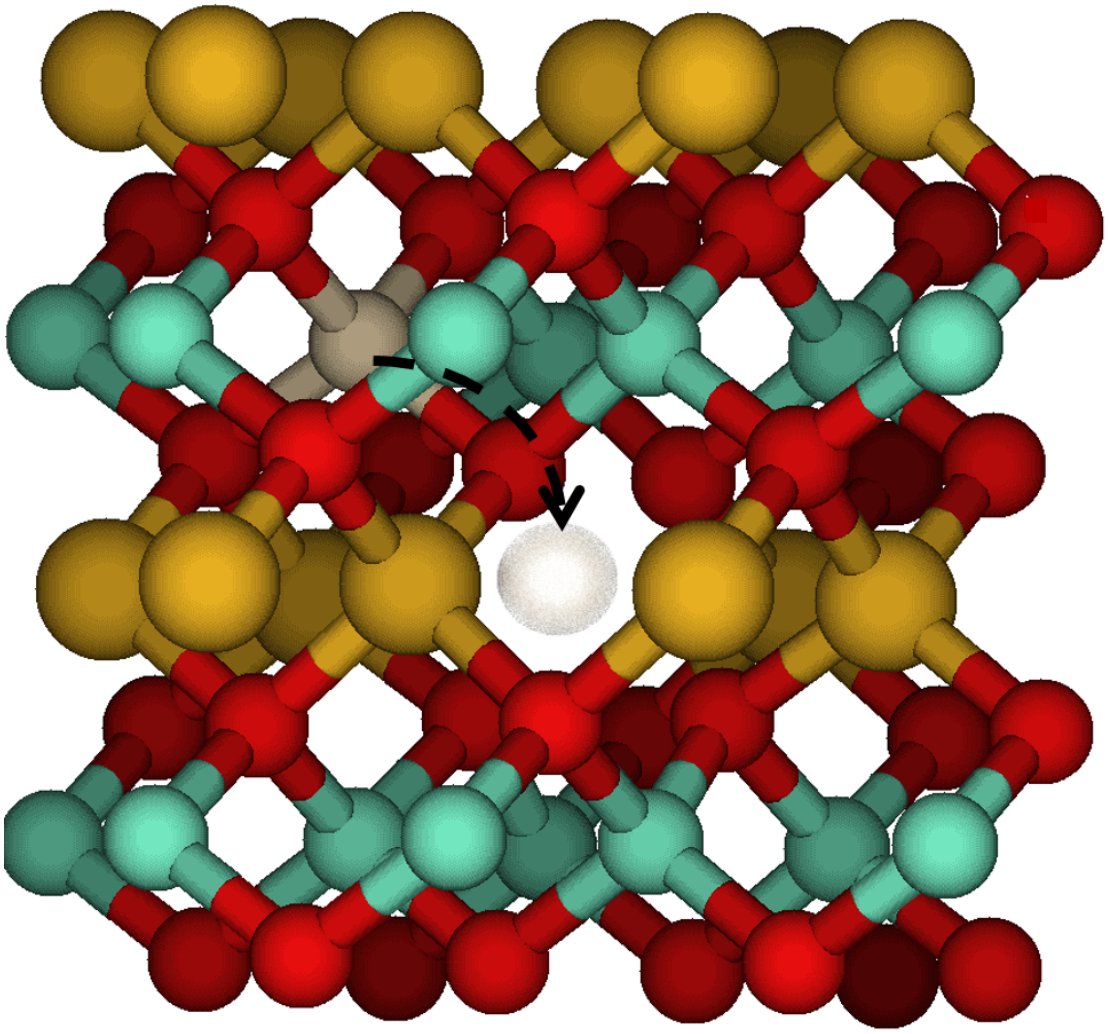


Figure 1.1: Example of a subsitutional dopant (shown in tan) and a vacancy defect in a 64-atom InGaAs crystal in a Zinc Blende structure and CuAu-I ordering. As, In and Ga are respectively shown in red, yellow and green. The black dashed arrow denotes a possible transition whereby the substitutional defect would move to the nearby vacancy site.

Dopant atoms can be grown-in or incorporated post-growth in the host semiconductor crystalline lattice [34]. One post-growth incorporation method called ion implantation [35] involves the acceleration of atoms to the surface of the crystal so they can be incorporated in the crystal. This is usually followed by one or more annealing steps at moderate to high temperatures to repair damages and promote the migration of impurity atoms to sites of interest to create n- or p-type doping. [36] This process of dopants migrating to sites of interest to become functional defects by providing additional charge carrier for conduction is called "activation." In practice, only a fraction of dopant atoms incorporated in a semiconductor will become activated. Therefore, a key technological and scientific question around doping is related to how one increases the percent activation of a certain dopant in a given semiconductor. Answering this question usually involves developing a good understanding of the thermodynamics and kinetics of those dopants in the host lattice. Such understanding is also important to prevent significant diffusion of those dopants during further device fabrication steps which could lead to dopant deactivation and potential device malfunctions.

1.4 Silicon as a dopant in III-Vs

Silicon is a common dopant in III-V materials. It has been used as an n-type dopant in InAs [37, 38], GaAs [39, 40], GaN [41, 42, 43], InGaAs [44, 45, 46] and so on. However, outstanding questions still remain regarding silicon's behavior as a dopant in III-Vs, and more specifically in the In-Ga-As family of compound semiconductors. In particular two fundamental scientific challenges with technological ramifications exist towards the full adoption of silicon as a dopant in

InGaAs: anomalous diffusion and compensation effects. Below we give more details about those two issues.

One of the earliest work on silicon diffusion in III-Vs was performed by Vieland [47] and Antell [48] who showed, in 1961 and 1965 respectively, that As overpressure affects Si diffusion in GaAs. In Vieland's work, the diffusion of various impurities in GaAs was studied as a function of ambient As pressure. At 1150 ^oC, Si diffusion showed dependence on As pressure up until a threshold of 1 atm above which diffusion became independent of pressure [47]. Starting with the assumption that arsenic pressure increases the concentration of Ga vacancies due to interaction between the As vapor and GaAs surface, they attributed the observed increase in Si diffusion with As pressure to a diffusion mechanism mediated by Ga vacancy. Within their framework, after the population of Ga vacancies became saturated at high As pressure, Si diffusion became independent of pressure as the Ga vacancy population remains constant.

In another piece of work, Kavanagh *et al.* [49] measured Si diffusion at polycrystalline-Si/GaAs interface and noticed significant diffusion of Si in GaAs. Interestingly, they noticed that the Secondary Ions Mass Spectroscopy (SIMS) profiles did not follow error function curves that would be expected from a constant diffusion coefficient. This suggested two different regimes for Si diffusion in GaAs: a concentration-independent regime which dominates at low Si concentration and a concentration-dependent one at high dopant concentration. In another study from Schubert [50], the concentration-independent diffusion coefficient was measured and the values obtained were two orders of magnitude lower than their concentration-dependent counterparts.

Moreover, Deppe *et al* [51] showed that Si diffusion in GaAs is Fermi level

dependent. More specifically, n-type doping enhances diffusion whereas p-type doping has an inhibitory effect. They believed this was an additional proof that vacancies play an important role in Si diffusion based on the assumption that n-type doping is expected to lead to more Ga vacancies. In addition, the influence of the Fermi level was clearly demonstrated suggesting that charged species might be involved in the mechanisms at play during Si diffusion in GaAs.

Another poorly understood aspect of silicon in the InGaAs family of semiconductors is the compensation effect. In Antell's work [48], Si films were deposited on p-type GaAs followed by 900 and 1000 ^oC anneals. Arsenic pressure dependence was observed as well as a saturation in dopant activation around $10^{18}cm^{-3}$ for Si concentrations up to $10^{20}cm^{-3}$ documenting the first reported Si dopant compensation in GaAs. To explain this observation, the argument of amphoteric behavior was used. According to the authors, since Si is an amphoteric dopant in GaAs, at low Si concentration Si atoms occupy Ga sites rendering the material n-type. At high concentration, it is believed that Si_{Ga} and Si_{As} coexist and lead to the observed electrical compensation [52].

Similar results have been obtained for Si diffusion in InGaAs. Silicon is one of the preferred dopants for n-type InGaAs and was historically believed to have negligible diffusion in InGaAs [53]. However, significant diffusion has been recently reported for Si concentrations above 4 x 10^{19} cm^{-3} [1]. Also, experiments indicate an activation limit of 1.7 x 10^{19} cm^{-3}, even for Si concentrations above equilibrium [54, 45, 46]. Higher carrier densities can only be achieved during metastable growth [44]. From an application standpoint, unfortunately, this limit is below the desired level required for integration into next-generation devices. Hence, understanding the origin of this phenomenon of Si diffusion at

high concentrations is critical.

Based on these early studies, the prevailing assumption regarding silicon diffusion in III-V, specially in GaAs and InGaAs is that the mechanism is Ga-vacancy mediated. Also, it was believed that compensation effects are due to the amphoteric nature of silicon as a dopant in III-Vs. Based on such assumptions, the enhanced diffusion observed at high Si concentration would be due to a concurrent increase in Ga vacancy concentration which would lead to higher Si diffusivity.

Such hypothesis has a couple of potential deficiencies: First, it only provides a thermodynamic explanation and ignores any kinetic effects arising from the diffusion rates of the different species involved. Second, an increase in the cation vacancy population will not necessarily translate into higher overall Si diffusion if the energy barriers for the associated transitions are too high. On the other hand, if the limited diffusivity of silicon at low concentration is only due to a small population of group III vacancies, one would expect the energy barriers for vacancy-assisted diffusion to properly reflect such an assumption.

1.5 Research objectives

With the renewed interest in incorporating III-Vs in next-generation transistors the questions of anomalous diffusion and compensation effects have come back to the forefront and require proper addressing if this technology is to move forward. Although as explained above, a number of publications have attempted to explain this behavior (with a special focus on GaAs), it remains an open question as the exact mechanisms responsible for this uncommon diffusive behav-

ior are still largely unknown. Beyond the technological interest for developing such understanding for improved device engineering capabilities, working on this problem also holds a scientific importance.

In this thesis, we address these open questions using computational techniques capable of probing the underlying atomistic mechanisms responsible for the observed diffusion and activation behavior of Si dopants in III-Vs. We start from the point of view that the two problems of diffusion and compensation are inseparable and influenced in various degrees by the interplay between the thermodynamic and kinetic characteristics of various point-defects. The primary objective is to uncover atomistic-level phenomena responsible for the uncommon diffusive behavior of Si in III-V materials. More specifically, we study the kinetics and thermodynamics of Si dopants and intrinsic defects in three representative III-Vs compounds: InAs, GaAs and InGaAs.

We also develop a novel Machine Learning-based approach to accelerate future studies aimed at exploring atomistic transitions in crystals. We report on the methods developed as well as custom software tools that we have created to facilitate the use of this new approach. We show superior computational efficiency with reasonable accuracy compared to more expensive methods commonly used to conduct such studies.

In particular, the scope of this thesis included the following aims:

- Explore different diffusion mechanisms for vacancies, interstitial and substitutional defects in GaAs, InAs and InGaAs and draw conclusions regarding most probable mechanisms
- Create continuum models to validate proposed mechanisms using appro-

priate comparisons with experimental data

- Develop accelerated and more cost-efficient method to allow for extended studies in other alloys and with other dopants

1.6 Thesis outline

The remainder of this thesis is divided as follow: in chapter 2, we describe the different methods used in this study including density functional theory, nudge elastic band and machine learning. In chapters 3 to 5, we present and discuss intrinsic and silicon diffusion in GaAs, InAs and InGaAs respectively. In chapter 6, we use the atomistic results to inform and develop a continuum model that describes Si and other amphoteric diffusion in III-Vs. In chapters 7 and 8, we present some of our work related to using artificial intelligence and machine learning to accelerate the process of finding transition pathways while reduce the computational cost of conducting such diffusion studies. Finally, in chapters 8 and 9 we discuss future work and present our final conclusions.

CHAPTER 2

METHODS

2.1 Overview

In this thesis, we use various computational methods to access different time and length scales. With a special focus on accuracy, we use first-principles Density Function Theory (DFT) calculations to find stable configurations and associated energies of various point defects in the III-V semiconductors of interest. We then use the Nudged Elastic Band (NEB) method to find minimum energy paths that govern diffusion of the various species. The coupling of DFT and NEB is a powerful approach to explore the energy landscape of the III-V compounds at the most fundamental level and in an accurate fashion. The results gathered from those studies are then used to build a continuum model that readily compares with experimental diffusion studies. To accelerate the process of computing diffusion rates, we develop a novel and unique approach based on advanced Machine Learning techniques. In our approach, Neural Network models are built and trained to predict atomic forces. In the following sections we provide succinct descriptions of each method.

2.2 Density Functional Theory

Density Functional Theory (DFT) is a method developed to solve Schroedinger's equation as given by Eq. 2.1 where E is the energy, Ψ the wavefunction and $\hat{H}$ the Hamiltonian operator for a system of M nuclei and N electrons. $\hat{H}$ is given

in atomic units by Eq. 2.2 and includes the kinetic energy and electrostatic interactions of all the particles in the system with r_{ij}, r_{iA} and R_{AB} respectively denoting electron-electron, electron-nucleus and nucleus-nucleus distances, ∇^2 the Laplace operator and Z_A the charge associated with nucleus A. Since it is based directly on quantum physics, DFT is one of the most accurate and powerful computational techniques available to study the property and structures of materials. It allows access to a variety of molecular properties including ground state energies, reaction paths, atomization energies, etc. while providing atomistic resolution. It has been successfully used for a lot of problems ranging from from simple atoms, molecules, solids, quantum and classical fluids, etc. Its accuracy comes at the price of extensive computational cost which limits the system size (number of atoms) that can practically be studied to a few hundreds at best.

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\nabla_A^2 + \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B=1}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.2}$$

It is essentially impossible to compute exact solutions of Shroedinger's equation for any realistic system. Some approximations are necessary and DFT uses a few of them. One of those is the Born-Oppenheimer approximation which assumes that the electrons are moving in the field of fixed nuclei. This reduces the kinetic energy of the nuclei to zero and their potential energy to a constant V_{ext}. Also, at the core of DFT is the concept of "electron density" defined as the probability of finding any of the N electrons within a volume element $d\vec{r}$. Another important concept which aided in the development of DFT is the vari-

ational principle for the ground state which states that the ground state energy is a functional of the number of electrons (and thus the electron density since $N = \int \rho(\vec{r}d\vec{r})$) and the nuclear potential V_{ext}. In simpler terms, there is a relationship between the energy of a system and its electron density. Even with these new concepts and assumptions, using Shroedinger's equation to solve for the properties of a molecular system remained elusive.

Two fundamental theorems proposed by Hohenberg-Kohn have made it possible to solve Shroedinger's equation. The formulation of these theorems also marked the creation of density functional theory. The first theorem demonstrates that the electron density uniquely determines the Hamiltonian operator and thus the external potential and consequently the full many particle ground state and all its associated properties. This means not only that there is a functional between the external potential and electron density, but that functional is unique (to within a constant). The proof follows by contradiction and can be consulted on the original paper [55]. In other words, if we know the electron density we can compute all the properties of the system including total energy, atomic structure, etc..

The second Hohenberg-Kohn theorem is very similar to the variational principle and states that the functional that delivers the ground state energy of the system, does so if and only if the input density is the true ground state density. In other words, any electron density that satisfies the boundary conditions on the system will lead to an upper bound on the ground state energy of the system, unless that electron density is the true ground state electron density itself. In practice, this means that we can try to minimize the energy of the system to continuously improve on a trial electron density and that way, solve for the true

ground state density.

With these two theorems, it has become theoretically possible to forego solving for the complex 3N-dimensional wavefunction (where N is the number of electrons in the system) and instead solve for the simpler electron density which is only a 3-dimensional function of x, y and z. In practice however, it is impossible to know the exact form of the functional that maps electron density and energy. In fact, the non-classical self-interaction correction, exchange and coulomb correlation of the kinetic energy of the electrons as well as any electron-electron interactions are completely unknown. Therefore, other approximations, usually based on the Kohn-Sham (KS) equations (Eq. 2.3), are necessary. Within the KS framework, the orbitals Ψ_i of a reference non-interacting system with the same electron density as the real interacting one are used. An additional term called the exchange-correlation potential V_{xc} is introduced to account for all the unknown contributions to the Hamiltonians. V_{ext} and V_H in Eq. 2.3 are respectively the external potential due to the presence of the nuclei and the Coulomb potential.

$$[-\frac{1}{2}\nabla^2 + V_{ext}(r) + V_H[n] + V_{xc}[n]]\Psi_i(r) = \epsilon_i\Psi_i(r) \tag{2.3}$$

The KS equations have reduced the intractable problem of N interacting particles to a simpler problem of N non-interacting particles moving in an effective potential which includes the exchange-correlation contribution. However, even with this simplification, the exact form of the exchange-correlation potential is still unknown, hence the need for approximations. In the simple Local Density Approximation (LDA), the exchange-correlation potential only depends on the electron density at each point. The LDA approach has been proven to work

fairly well for a number of problems. However notable limitations such as underestimation of the bandgap energy of semiconductors and incorrect predictions of unstable configurations exist. Other approximations such as the Generalized Gradient Approximation (GGA) improve upon LDA by including the gradient of the electron density to account for non-homogeneity. In the meta-GGA approach, the second-order derivatives of the electron density are also included for increased accuracy. Other approximations such as hybrid functionals which include additional corrections have also been developed.

2.3 Nudged Elastic Band

The Nudged Elastic Band (NEB) [56, 57, 58, 59] method is one the most popular computational techniques available to find the Minimum Energy Path (MEP) between two stable configurations of a molecular system. The MEP gives access to the rate at which interconversion between those two states take place. The two configurations, usually referred to as initial and final states in the case of diffusion, or reactant and products in the case of reactions have to be known prior to using this approach. The NEB method has been successfully used in a lot of problems including diffusion in bulk crystals, surface diffusion and a number of chemical reactions, etc. [58, 56].

The NEB approach relies heavily on the theoretical framework provided by the Transition State Theory (TST) in which chemical reactions and diffusion events are assumed to take place in a potential energy landscape and an intermediate state called saddle point or transition state exists between the initial and final configurations (see Fig. 2.1). Within TST, thermal atomic vibrations

lead to collision events between atoms in the system. From time to time, some of those collisions will lead to the formation of the transition state and subsequently to the final state. The number of collisions that occur per unit time is called attempt frequency and only a fraction of those, determined by the energy difference between the initial state and saddle point, will successfully lead to the final state. The transition state corresponds therefore to a maximum in energy along the reaction path and the difference between its energy and that of the initial state is usually called energy barrier. The rate of the process is then given by an Arrhenius like expression shown in Eq. 2.4 where $E^{\dagger}$ is the energy of the transition state, E^{init} the energy of the initial state, k_B Boltzmann constant, T the temperature and A is called the pre-exponential factor which incorporates the attempt frequency.

$$k^{TST} = A \exp(-(E^{\dagger} - E^{init})/k_B T) \tag{2.4}$$

In NEB, intermediate configurations, known as "images," are chosen and maintained together with spring-like forces (hence the name "elastic band") acting only along the path of the reaction trajectory. Each image is allowed to relax only in the direction perpendicular to the reaction path to prevent images from falling back to the initial and final configurations. The problem formulated in this manner is then optimized to minimize the forces. A good overview of how to carry such an optimization is given by Herbol *et al.* [60]. A variant of this method called Climbing Image NEB moves one of the images uphill in the energy landscape along the reaction path to ensure the saddle point is identified [57].

In this thesis, all the forces and energies are calculated using DFT to ensure

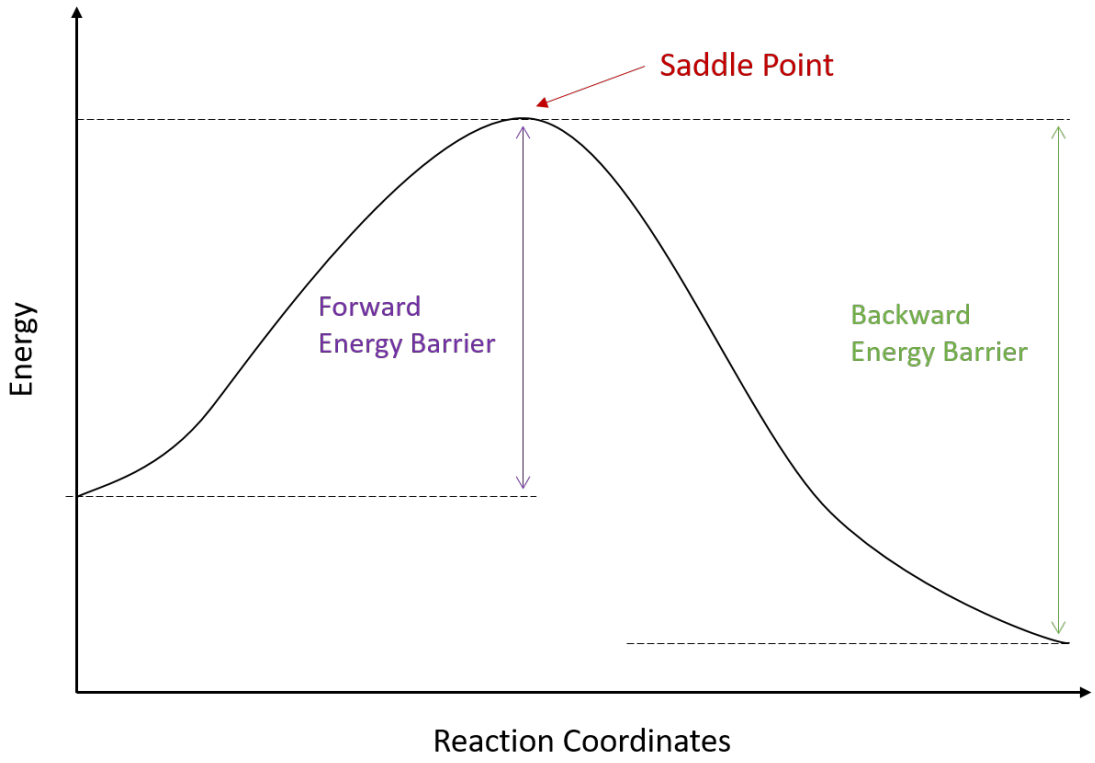


Figure 2.1: Example energy profile along an atomistic jump or chemical reaction. The energy barriers give access to the forward and backward rates.

accurate results. Initial and final configurations are identified using the symmetry of the crystals of interest. Their geometries are then optimized to ensure they are stable and sit at the bottom of a well in the energy landscape. Simple linear interpolation between those two configurations are then used to create images for the NEB calculation. After initial convergence of the MEP, climbing-image NEB is used to ensure the saddle point is found. In case more than one saddle points are found along a given diffusion path, the trajectory is broken down into individual steps to ensure there is only one transition state per diffusion path. This process is repeated for a number of point-defects in GaAs, InAs and InGaAs to build a library of energy barriers corresponding to individual diffusion events and thereby gain insights into favorable diffusion mechanisms.

2.4 Machine Learning

Machine Learning (ML) is a subfield of computer science and artificial intelligence that focuses on building algorithms that can learn from and make predictions based on data. Recent developments in (ML) algorithms have afforded significant progress in diverse application areas, including medical diagnosis, [61, 62, 63] speech recognition, [64, 65, 66] face recognition,[67, 68] the discovery of new planets, [69] and many others. Following such profound successes in those areas, researchers in materials science are increasingly interested in investigating the promise of similar techniques for the prediction of material properties, [70, 71, 72] the discovery of new materials,[73, 74, 75] force field development[76, 77, 78, 79] for use in molecular dynamics simulations, and so on. In these applications, it has been shown that data from both experiments and simulations can be leveraged to make predictions of a large number of

properties including gas uptake capacity in Metal Organic Frameworks, [71] crystal structure predictions,[80] band gap energies of double perovskites,[81] heat capacity of organic molecules, [70] *etc*.

The primary goal of all machine learning techniques is to learn from a given dataset to create models that can generalize well and make accurate predictions on unseen data. There are two general classes of machine learning setup: supervised and unsupervised. In the former, all data points are provided with a corresponding "label" or "response" whereas in the latter the data is unlabeled and the algorithm is tasked with determining its inherent structure. In supervised learning, a given problem is a regression task if the output or labels are continuous, otherwise it is a classification problem. More formally, in a supervised ML setup, labeled training data points $(\vec{x}, y)$ where $\vec{x} \in R^d$ is the input, y is the label or output and d the dimensionality of the problem, are provided. The entire dataset D of n data points can be represented as $D = (\vec{x}_1, y_1), \ldots, (\vec{x}_n, y_n) \subseteq \mathcal{R}^d \times \mathcal{C}$ where $\mathcal{C}$ is the label space. The goal of the ML algorithm is to find a function h: $\mathcal{R}^d \rightarrow \mathcal{C}$ such that $h(\vec{x}_i) \approx y_i$ for all $(\vec{x}_i, y_i) \in D$ (training) and also $h(\vec{x}_i) \approx y_i$ for all $(\vec{x}_i, y_i) \notin D$ (testing). It is important to note that the choice of a specific ML algorithm inherently encodes some assumptions regarding the type of function h that is sought and consequently the type of problem that is being solved. This is usually called the "hypothesis set" represented by $\mathcal{H}$. The free lunch theorem states that such assumptions are required for any successful ML algorithm. In other words, there is no single ML algorithm that will work well for all problems.

In order to find the best function h for a dataset D, a suitable loss function $\mathcal{L}$ is defined to quantify how well a specific solution works. Example loss func-

tions include the squared loss and absolute loss functions given respectively by Eq. 2.5 and 2.6. The right machine learning model h is therefore the function that minimizes the loss, *i.e.* $h = argmin_{h \in \mathcal{H}} \mathcal{L}(h)$. To ensure proper generalization of a given learning task, the dataset is usually divided into a training, validation and testing set. h is then minimized based on the error on the validation test while training is conducted on the training set. The testing set is left for evaluation of the generalization error that would be made on completely unseen data points after training is complete. In many algorithms, the function h involves a vector of weights $\vec{w}$ which has to be optimized such that the loss function is minimized. The problem is thus an optimization task formalized as follow: $min_{\vec{w}} \frac{1}{n} \sum_{i=1}^{n} l(h_{\vec{w}}(\vec{x_i}), y_i) + \lambda r(w)$ where $\lambda r(w)$ is called a "regularizer" and is added to penalize complexity and avoid overfitting to the training dataset.

$$\mathcal{L}_(h) = frac1n \sum_{i=1}^{n} (h(\vec{x_i}) - y_i)^2 \tag{2.5}$$

$$\mathcal{L}_(h) = frac1n \sum_{i=1}^{n} |h(\vec{x_i}) - y_i| \tag{2.6}$$

A large number of ML algorithms have been developed over the years and open-source packages such as TensorFlow, SciKit-Learn [82, 83] exist with standard implementations. A comprehensive review of all the ML algorithms is outside the scope of this thesis but the reader is referred to the following publications for more details. In this work we primarily use Artificial Neural Networks which are described in detail in the next section (Section 2.5).

A number of practical considerations are necessary to ensure machine learning studies are successful. For instance, if the dataset is not representative of the

input space, the learned model will be highly dependent on the data included in the training process leading to an increase in variance of the learned model. Since this problem can usually be solved by simply increasing the size of the training dataset, it is necessary to pay particular attention to ensure the training dataset has a large enough number of points. However, in some situations, the method used to generate or collect the data might simply not cover enough of the sampling space, expert knowledge of the application domain and advance data generation techniques are usually necessary to solve this type problem. In addition, the choice of the right algorithm for a given learning task is usually not as straightforward as one might think and constitutes an active area of research. It is also worth noting that a key step in all machine learning studies is the pre-processing of the training dataset in a way that is suitable for learning. This aspect is treated in detail in the chapter 7 on molecular fingerprints.

2.5 Artificial Neural Networks

Artificial Neural Networks (ANN) are specific types of machine learning algorithms vaguely inspired by how the human brain works. ANNs are comprised of neurons and connections used to transfer signals (usually a numeric value) between them. ANNs are usually organized into layers of neurons whereby each layer performs a linear combination over inputs and pass the results along to the next layer. At the end of the process, the last layer, which usually has only one neuron, uses an activation function to return the final result.

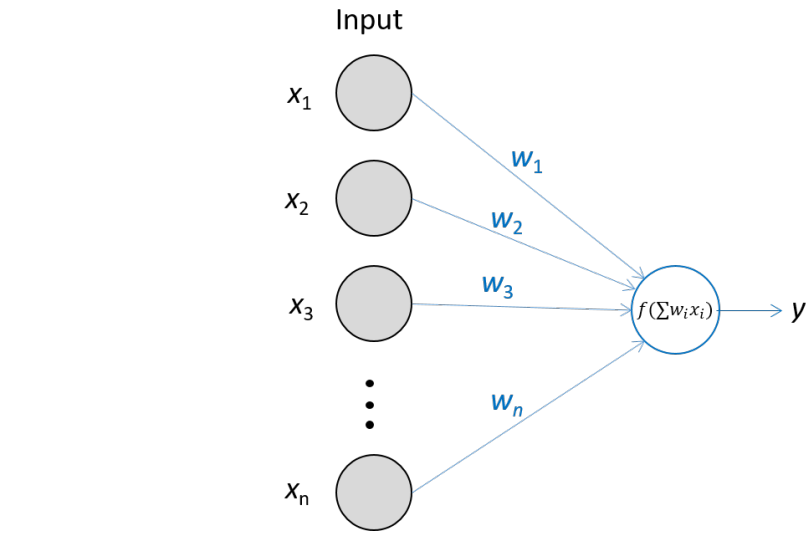
Historically, ANNs can trace their origin back to the late 1950s with the invention of the Perceptron at Cornell University by Frank Rosenblatt. As

illustrated in Fig. 2.2 the Perceptron works by taking linear combination $\sum_{i=1}^{m} w_i x_i + b$ of inputs (real-valued vector) $\vec{x}$ and applying a bias b. The result is then fed to a Heaviside step function to produce the final result, which in this case is a binary value. The Heaviside step function (given by Eq. 2.5) is in this case the "activation function" and the perceptron which can be seen as a single layer ANN. During training, the weights w_i are optimized and the final trained Perceptron model can be used as a linear classifier.

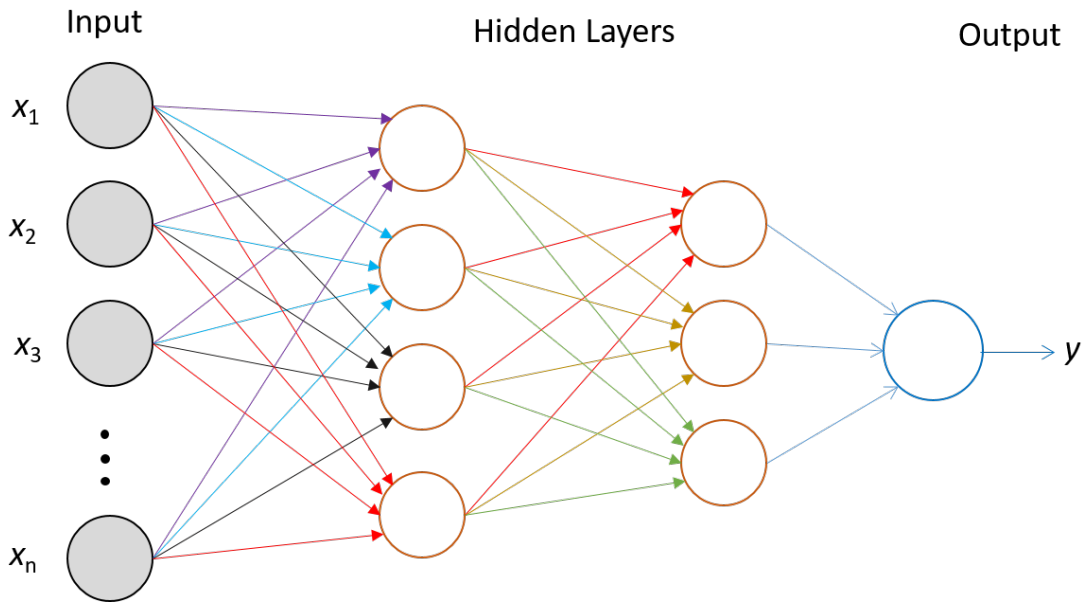
$$f(x) = \begin{cases} 1, & \text{if } \sum_{i=0}^{m} w_i x_i + b > 0 \\ 0, & \text{otherwise} \end{cases}$$

ANNs are essentially an extension of this idea whereby multiple perceptrons are used both simultaneously and consecutively (see Fig. 2.2). Hence, ANNs are sometimes referred to as Multi-Layer Perceptrons (MLPs). Until the 1970s, however, MLPs were very limited in size because of challenges related to the optimization of the weights. In 1982, a method called "backpropagation" was successfully applied to train Neural Networks in a way that allows the use of bigger networks. Additional advancements in computing power, parallel processing and algorithms over the past 20 years have greatly improved our ability to train ANNs to solve increasingly challenging problems using bigger and more complex architectures. Those new developments have lead to the use of deep learning (which simply refers to the use of ANNs with many layers), Recurrent neural netowrks (RNNs), Convolutional Neural Networks (CNNs), etc. These ANNs have led to much success in natural language processing, image and face recognition, medical diagnosis, etc. [64, 65, 66, 69, 67, 68].

In this thesis, ANNs are used to learn from data on molecular configurations,



(a)



(b)

Figure 2.2: (a) Schematic representation of a the Perceptron. This basic implementation of a Neural Network has only one neuron and uses a Heaviside step function to transform a linear combination of the features of the input vector into a label(0 or 1) that can be used for binary classification. (b) Schematic representation of an artificial Neural Network. This can be seen as an extension of the Perceptron whereby multiple neurons are combined into layers allowing to learn more complicated functions.

atomic forces and energies generated from DFT calculations. Part of those DFT calculations were performed to calculate the energy barriers of various diffusion events. Those computations are very expensive and involve a significant number of data generated during the course of the calculation that are never used otherwise, leading to waste of computation. For example, a number of configurations will be explored during geometry optimization as well as during the process of finding the minimum energy path for a given diffusion mechanism but since only the final result is of interest, those data generated during intermediate steps are discarded. Here, we explore how we can use ML to learn from those data and make predictions for unseen scenarios. This novel approach can take advantage of data previously generated while limiting the amount of new data that needs to be generated.

CHAPTER 3

INTRINSIC AND SILICON DIFFUSION IN GaAs

3.1 Introduction

GaAs is one of the most studied III-V compounds. For example, a quick search on the SCOPUS database show that GaAs has been studied at least twice as much as InAs and InGaAs (see Fig. 3.1). As shown in Figure 3.2, GaAs adopts a Zinc Blende structure with tetrahedral coordination similar to the diamond cubic structure of silicon. The only difference is the presence of separate As and Ga sublattices (hence the name Zinc Blende instead of diamond cubic for the structure). As with other III-Vs, it has a direct bandgap and high electron mobility which make it very attractive for applications in high-speed electronics, photo-electrical energy conversion, detectors, etc.

With such a large body of work available on GaAs, it has been traditionally used as a model system to understand diffusion in InGaAs. Mainly, reports dating back to the 1980s have proposed different models attempting to explain existing data on Si diffusion in GaAs. Corresponding analytical expressions were also developed to provide quantitative and predictive capabilities describing diffusion in this specific III-V material. Recent studies on InGaAs have relied on those previous models to try to explain the newly observed concentration-dependence of Si diffusion in InGaAs. However, multiple issues exist with this approach. Some of those issues were partly addressed in Section 1.4. In this chapter, we will present an overview of existing experimental data on Si diffusion in GaAs as well as some data on diffusion of other group IV elements. This is followed by a brief overview of the two main models for diffusion in GaAs.

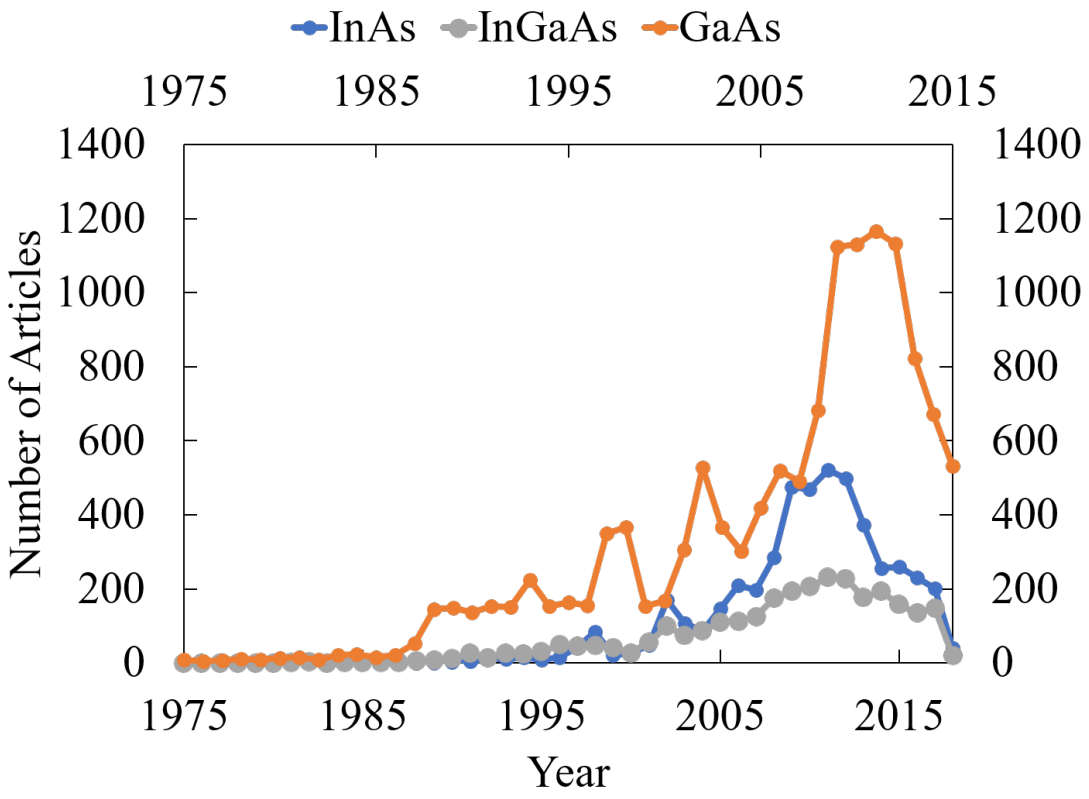


Figure 3.1: Number of articles published between 1975 and 2018 with InGaAs, GaAs or InAs in their title or list of keywords (data from SCOPUS - www.scopus.com).

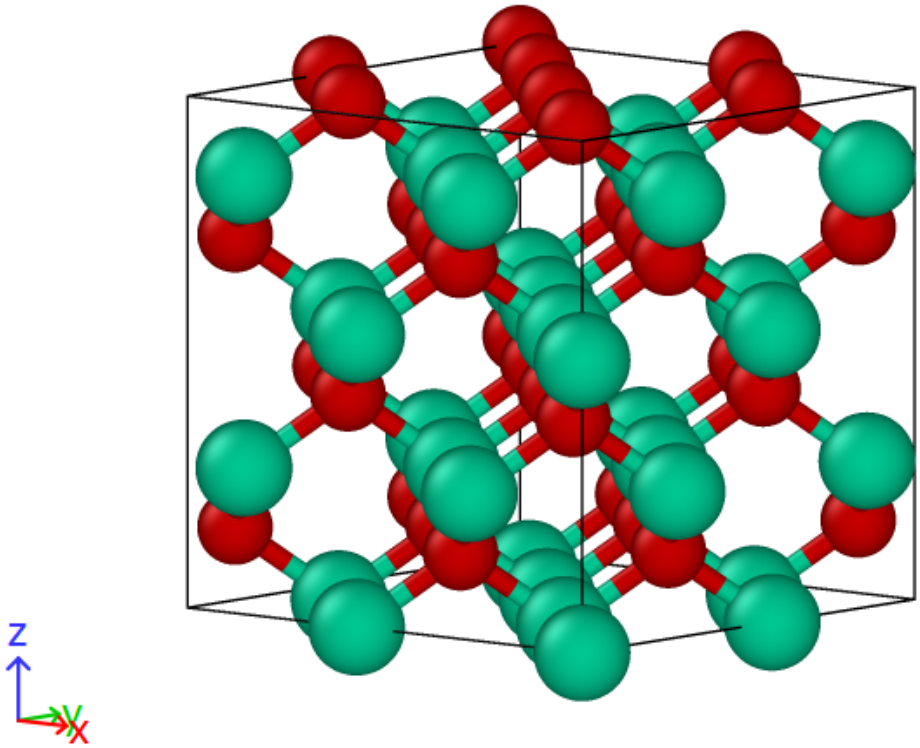


Figure 3.2: Example of a 64-atom GaAs crystal in the Zinc Blende structure with As shown in red and Ga in green.

We will then present additional NEB calculations that were performed to clarify some of the assumptions and finally, we will discuss their meaning.

3.2 Diffusion results from Literature

3.2.1 Silicon

As mentioned in section 1.4, multiple studies have shown that Si diffusion in GaAs is heavily concentration-dependent [47, 50]. Results have also shown that there are compensation effects in heavily doped GaAs, whereby increase in Si concentration beyond a certain dopant level does not lead to an increase in carrier concentration.

A host of authors have studied silicon diffusion in GaAs. The list include Ahlgren *et al.* [84] who used secondary-ion mass spectrometry (SIMS), nuclear resonance broadening techniques and Rutherford backscattering spectrometry channeling techniques to confirm the concentration-dependence of Si diffusion in GaAs. Other studies by Murray *et al.* [85] and Keller *et al.* [40] have reproduced the concentration-dependent diffusion of Si in GaAs. An enhanced diffusion with doping level (suggesting a Fermi level dependence) and As overpressure was observed leading the authors to believe that a vacancy-mediated diffusion mechanism is the probable diffusion mechanism.

Using DFT calculations, Zhang and Northrup [86] showed for the first time in 1991 that Ga vacancies V_{Ga} have low formation energy, specially in As-rich growth conditions, and prefer to adopt a charge state of -3. According to

the authors, this result confirmed that V_{Ga} must play a role in the diffusion of silicon in GaAs. In 1994, Daþrowski and Northrup [87] followed up by proposing an interstitial-vacancy mechanism for Si diffusion in GaAs whereby a $V_{Ga} - Si_{int} - V_{Ga}$ configurations would mediate the Si jump from a Si_{Ga} substitutional configurations to a V_{Ga} Ga vacancy sites.

3.2.2 Other group IV dopants: Tin and Germanium

In general, Sn is reported to have slow diffusion in GaAs with diffusivity values around 10^{-12} at 1000 ^oC. Data published by Tuck and Badawai [88] on Sn diffusion over the temperature range of 850 ^oC - 1100 ^oC showed that Sn diffusion was different when the GaAs sample was n-doped versus undoped. At low temperatures, diffusion in the n-doped sample was at least one order of magnitude higher than the undoped sample. Whereas at high temperature there was no difference in the two samples. Tuck and Badawai developed a diffusion model which attributed this behavior to the presence of a higher electron concentration in the n-doped sample. At high enough temperatures, thermally generated carriers in the undoped sample which increase the electron concentraion would then explain the similarity between the diffusion in the two smaples at high temperature. Essentially, these results and associated argument suggest that the diffusion coefficient strongly depends on electron concentration. The proposed mechanism was that most of the Sn exists in the form of Sn^+_{Ga} and diffuses via gallium vacancies. Additionally, an equilibrium assumption between neutral and singly-charged vacancy was made ($V^0_{Ga} + e^- \leftrightarrow V^-_{Ga}$) leading to an expression for the charged vacancy concentration of the form $[V^-_{Ga}] = Kn[V^0_{Ga}]$ where n is the electron concentration and K the equilibrium constant. Now assuming a

constant $[]V_{Ga}^{0}]$ at constant temperature, Tuck and Badawi found that $[V_{Ga}^{-}] \propto n$ which would explain why an increase in electron concentration would give rise to an increase in Ga vacancy concentration and hence diffusion.

An alternative model was later proposed by Shaw (1984) [89] who suggested that instead of Sn_{Ga}^{+} being the mobile species, a complex of the form $V_{Ga} - Sn_{As}^{-} - V_{Ga}$ was the primary diffusing species. The Sn_{Ga}^{+} species were considered immobile. This new model assumes that vacancies remain neutral and that there is an interconversion between the complex $V_{Ga} - Sn_{As}^{-} - V_{Ga}$ and Sn_{Ga}^{+} involving two electrons. These new assumptions along with the assumption of gallium-rich conditions lead to a diffusion coefficient which varies with the squared of Sn_{Ga}^{+} concentration. The reason used by Shaw to suggest this alternative model was the proposed existence of a $V_{Ga}Sn_{As}^{-}V_{Ga}$ based on studies done by Panish (1973) [90].

It is important to note here that the experimental results as well as the ensuing arguments used to explain Sn diffusion in GaAs are similar to the situation for Si diffusion in GaAs where different hypothesis and assumptions lead to conflicting models to explain available experimental data. Although not discussed here, Tuck and Badawai also suggested that compensation effects might exist whereby the electron concentration remains lower than the Sn concentration at high dopant concentration. As with silicon diffusion in GaAs, in general, there is little evidence presented to support assumptions of most stable species, charged state and favorable diffusion mechanism.

Results for Germanium diffusion in GaAs are not very different from results for Sn and Si. They show diffusion on the order of 10^{-14} at 1000 ^{o}C, two orders of magnitude smaller than for Sn. Results point to the fact that diffusion takes

place on the III sublattice and $Ge_{Ga}V_{Ga}$ complexes form faster than Ge_{As} and could also be reponsible for high compensation effects that were observed.

3.2.3 Overview of existing analytical diffusion models

There are two main models that attempt to explain the existing data on Si diffusion in GaAs. The first one was developped in 1985 by Greiner and Gibbons [91] and is based on the main assumption that the $Si_{Ga} - Si_{As}$ pairs are fast diffusers whereas the Si_{Ga} and Si_{As} individual species are slow diffusers. Other assumptions of this model include that at high dopant concentration, Si in GaAs is heavily compensated meaning the concentration of donors is the same as the concentration of acceptors ($C_d = C_a$) and there is equilibrium between neutral pairs and their corresponding donor and acceptor single atoms as described by the equation below.

$$\mathrm{Si_{Ga}}^{+1} + \mathrm{Si_{As}}^{-1} \rightleftharpoons \mathrm{Si_{Ga}}{-}\mathrm{Si_{As}} \tag{3.1}$$

This model has led to an expression for the effective concentration-dependent diffusivity of the form given by the equation below:

$$D_{eff} = [1 - (1 + 2C_t/K_p)^{-1/2}] \tag{3.2}$$

Another model was proposed by Yu *et al.*[52] in 1989 and make vastly different assumptions compared to the Greiner and Gibbons model. Rather than

assuming that $Si_{Ga} - Si_{As}$ pairs are fast diffusers, Yu *et al.* starts with the assumption that $Si_{Ga} - Si_{As}$ are immobile while Si_{Ga} and Si_{As} are mobile species. Yu's model also takes into account changeover reactions that transforms Si_{Ga}^{+1} to Si_{As}^{-1} with the involvement of a As or Ga vacancies. The compensation reaction is still the same as in Greiner and Gibbons. These new assumptions have led to a rather complicated expression for the concentration-dependent diffusivity of Si in GaAs with multiple fitting parameters. However, this model has been shown to fit a wide range of different experimental data and also has the merit of taking the Fermi level into account. Hence, it has been historically favored over Greiner's model and remains one of the main reasons cited by authors who mention a vacancy-assisted mechanism as the primary diffusion mechanism for Si in GaAs.

3.3 Methods

To evaluate those two models and investigate the atomistic origin of the concentration-dependence of silicon diffusion in GaAs, we perform Nudge Elastic Band calculations at the DFT level. We found that a cut-off energy of 75 Ry and a $3 \times 3 \times 3$ k-mesh achieved a targeted convergence of 0.005 eV or less in DFT energy. The simulations were performed with norm-conserving Perdew-Zunger pseudo-potentials within the Local Density Approximation (LDA) level of theory. An LDA level of theory has been shown to be an appropriate choice for defect formation and migration energy calculations for III-V systems of this type [92, 93]. This has also been confirmed by some of our calculations (not reported here) that show negligible improvement when using GGA rather than LDA for Nudged Elastic Band (NEB) calculations in the closely related system

of CuAu-I $In_{0.5}Ga_{0.5}As$.

We start by optimizing a 64-atom GaAs supercell as shown in Fig. 3.2. The equilibrium lattice parameters were found to be 5.53 Å in reasonable agreement with experimental value of 5.65 Å. We then introduce defects by removing one atom to create a vacancy and replacing one As or Ga atom with silicon to create a substitutional defect. This step is followed by geometry optimization of the defect-containing supercells wto allow the atoms to relax and reach their equilibrium positions.

We then perform NEB calculations between a suitable set of initial and final configurations. Five 5 intermediate configurations (or images) are used along the reaction coordinate to determine the minimum energy path of each transition. The vacancy-assisted atomic jumps studied in this manner are given in Table 3.1 and include first nearest neighbor jumps between Ga and As sites which was suggested by Yu *et al.* as important. They also include second nearest neighbor jumps involving single Si atoms $Si_{Ga} - SiAs$ pairs diffusing on the As or Ga sublattice.

Although not present in the mechanisms proposed in literature and reviewed above, we also perform studies for Si interstitial species and associated jumps. In fact, arsenic split interstitial configurations have been recently identified in GaAs and since arsenic is a known fast diffuser in GaAs, we also investigate if such split intersitial configurations could also exist for Si and if so, what are the energy barriers involved.

Vacancy-Assisted Transitions			
	Atom Moving	Initial Lattice Site	Final Lattice Site
First Nearest Neighbor Jumps	Si	Ga As	As Ga
Second Nearest Neighbor Jumps	Si	Ga As	Ga As
Second Nearest Neighbor Pair Jumps	$Si_{Ga} -$ Si_{As}	Ga As	Ga As

Table 3.1: List of vacancy-assisted transitions studied in this work.

3.4 Results and discussion

3.4.1 Vacancy-assisted diffusion

Results for the change-over mechanism that transforms Si_{Ga} into Si_{As} and vice versa are shown in Fig. 3.3. The corresponding transitions are described in Equation 3.3 and clearly show that this jump is possible with a very low energy barrier. This result also show that Si only marginally prefers to stay on the Ga site compared with the As site when there is a vacancy present within one nearest neighbor. In other words, the $Si_{Ga}-V_{As}$ has similar stability as the $Si_{As}-V_{Ga}$. Consequently, with this change-over transition favorable, Si diffusion does not need to be constrained in a given sublattice. It can potentially proceed on either the As or Ga sublattices with successive change-over jumps. This means that a good diffusion model should also take this type of transition into account, as did the Yu *et al.* model [52].

As shown in Fig. 3.4, vacancy-assisted second nearest neighbor jumps of Si_{Ga} and Si_{As} individual species have energy barriers over 2 eV. Compared to the change-over transition, second nearest neighbor jumps of individual Si atoms in substitutional positions, though possible, are less favorable. These results also show that jumps on the Ga sublattice involve a slightly lower energy barrier of 2.1 eV compared to 2.5 eV for jumps on the As sublattice.

Lastly, we evaluated Greiner and Gibbons assumption that the mobile species are $Si_{Ga} - Si_{As}$ pairs. We perform calculations with a $Si_{Ga} - Si_{As}$ pair with either a V_{Ga} or V_{As} at a second nearest neighbor distance. The NEB transition then involve one of the Si atoms moving to the nearby vacant site. As

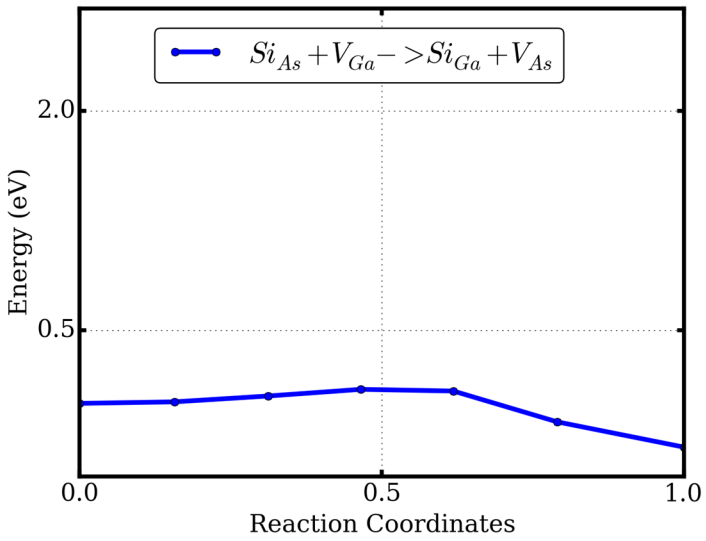


Figure 3.3: Si diffusion by first nearest neighbor vacancy-assisted jumps.

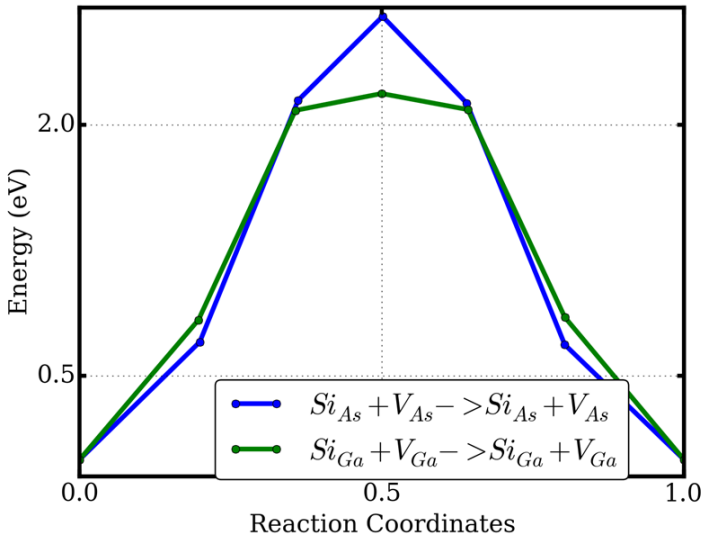


Figure 3.4: Energy barriers for Si diffusion by second nearest-neighbor vacancy-assisted jumps.

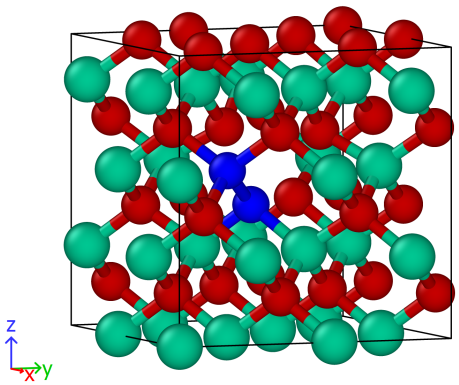
shown in Fig. 3.6, the movement of Si pairs in this manner involve higher energy barriers compared to the change-over transition, but lower compared to the corresponding jump by single Si species. In other words, $Si_{Ga} - Si_{As}$ pairs are faster diffusers than Si_{Ga} and Si_{As} defects. This means that, although Si pairs do not benefit from significantly lower energy barriers, Greiner and Gibbons were partially correct in assuming that Si pairs are more mobile than individual Si atoms.

Overall, results presented here show that the various transitions happen in the following order from more favorable to less favorable: change-over jumps, Si pairs second nearest neighbor jumps on the Ga sublattice, individual Si jump on the Ga sublattice, Si pairs second nearest neighbor jumps on the As sublattice and finally, individual Si jump on the As sublattice. These transitions are also given in Equations 3.3 to 3.7. These NEB results essentially show that neither the individual Si_{Ga} and Si_{As} atoms nor the $Si_{Ga} - Si_{As}$ pairs would qualify as "mobile" species through vacancy-assisted jumps and therefore, both Greiner and Yu's models might be wrong to some extent.

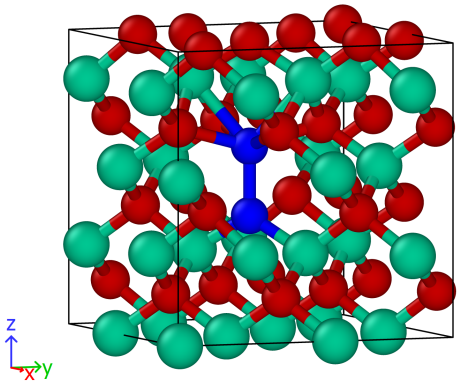
$$\mathrm{Si_{As}} + \mathrm{V_{Ga}} \underset{0.4\ \mathrm{eV}}{\overset{0.1\ \mathrm{eV}}{\rightleftharpoons}} \mathrm{Si_{Ga}} + \mathrm{V_{As}} \tag{3.3}$$

$$\mathrm{Si_{Ga}{-}Si_{As}} + \mathrm{V_{Ga}} \underset{1.84\ \mathrm{eV}}{\overset{1.84\ \mathrm{eV}}{\rightleftharpoons}} \mathrm{Si_{Ga}{-}Si_{As}} + \mathrm{V_{Ga}} \tag{3.4}$$

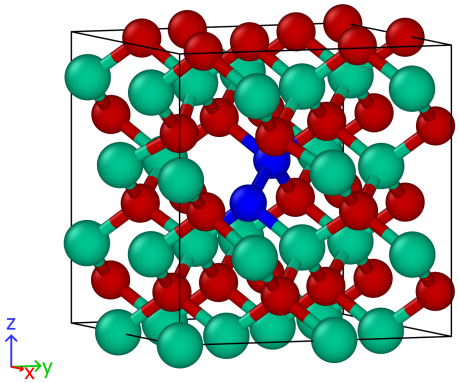
$$\mathrm{Si_{Ga}} + \mathrm{V_{Ga}} \underset{2.19\ \mathrm{eV}}{\overset{2.19\ \mathrm{eV}}{\rightleftharpoons}} \mathrm{Si_{Ga}} + \mathrm{V_{Ga}} \tag{3.5}$$



(a) Initial configuration



(b) Saddle point



(c) Final configuration

Figure 3.5: $Si_{Ga} - Si_{As}$ pair undergoing a vacancy-assisted transition on the Ga sublattice in a 64-atom GaAs supercell. Ga, In and Si are respectively shown in green, red and blue.

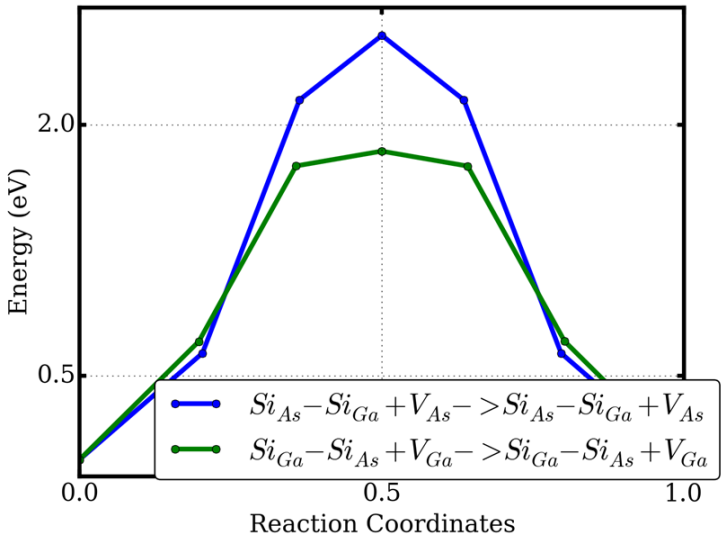


Figure 3.6: Energy barriers for Si-Si pairs diffusing by second nearest-neighbor vacancy-assisted jumps.

$$\mathrm{Si_{Ga}}-\mathrm{Si_{As}} + \mathrm{V_{As}} \underset{2.53\ \mathrm{eV}}{\overset{2.53\ \mathrm{eV}}{\rightleftharpoons}} \mathrm{Si_{Ga}}-\mathrm{Si_{As}} + \mathrm{V_{As}} \tag{3.6}$$

$$\mathrm{Si_{As}} + \mathrm{V_{As}} \underset{2.65\ \mathrm{eV}}{\overset{2.65\ \mathrm{eV}}{\rightleftharpoons}} \mathrm{Si_{As}} + \mathrm{V_{As}} \tag{3.7}$$

3.4.2 Interstitial impurities

Recent results have shown evidence of the presence of As atoms in stable split interstitial configurations in GaAs [94]. More recent studies have shown that those split-interstitial species are very fast diffusers with energy barriers on the order of 0.05 eV.[**?**]. Here we confirmed that such arsenic split interstitial configurations do exist in GaAs as shown in Fig. 3.7.

Knowing that As is a fast diffuser in GaAs, we investigated whether Si split interstitial configurations would be stable. As shown in Fig. 3.8, this is indeed the case. Interstitial silicon atoms can exist in split interstitial configurations whereby they share one lattice site with an As atom.

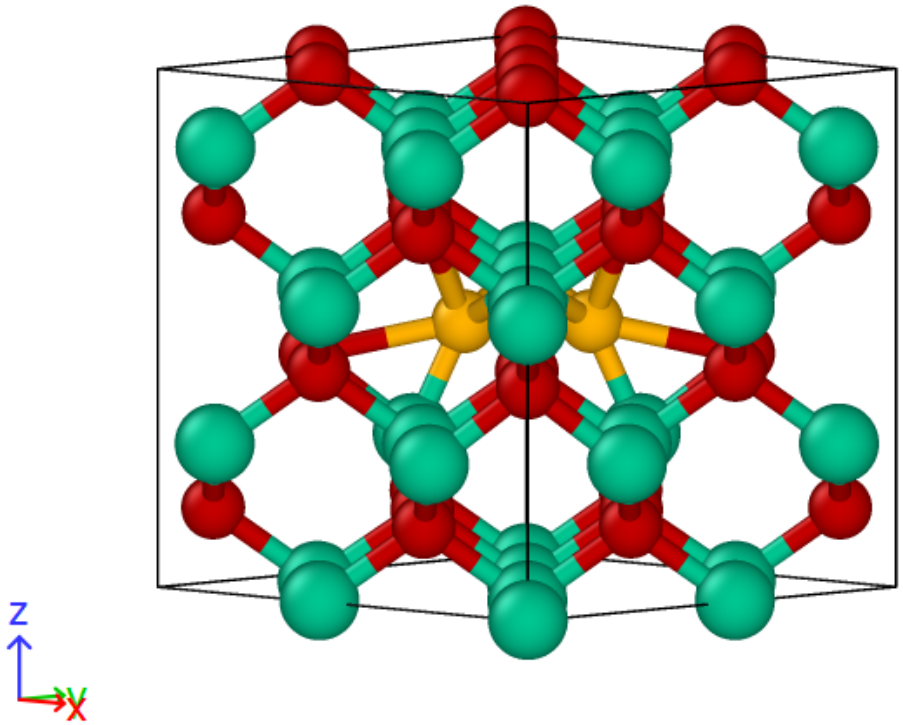


Figure 3.7: Optimized geometry for a As atom (in yellow) in a split interstitial configuration in GaAs. Ga atoms are in green and As atoms in red.

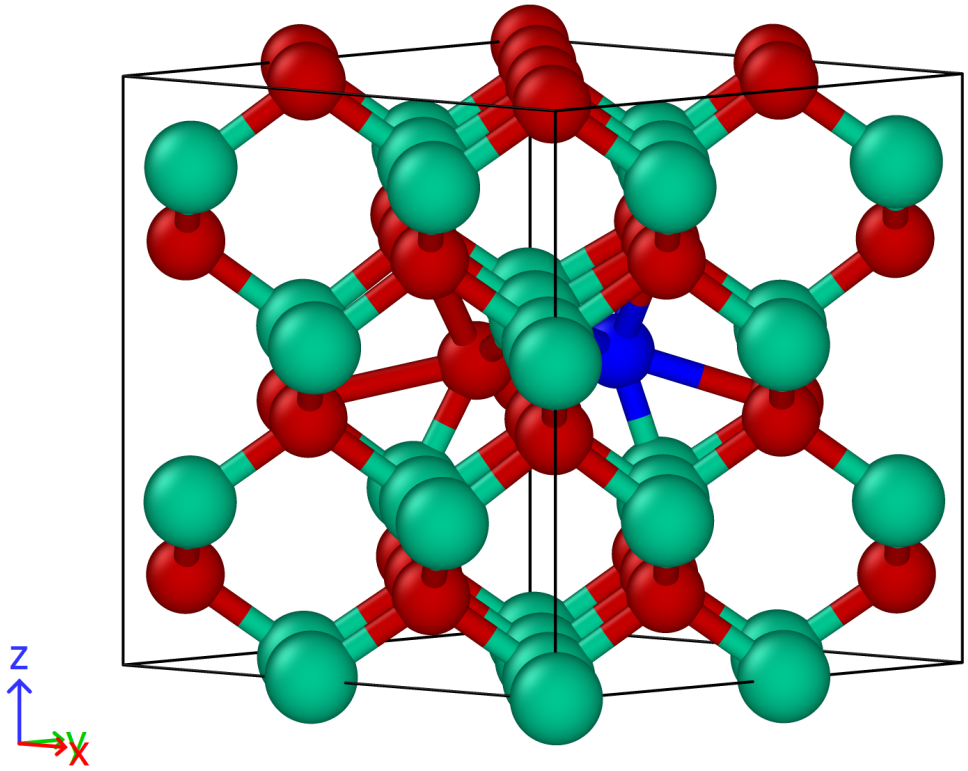


Figure 3.8: Optimized geometry for a Si atom (in blue) in a split interstitial configuration in GaAs. Ga atoms are in green and As atoms in red.

3.5 Conclusions

In contrast to InAs and InGaAs (the ternary alloy made of GaAs and InAs), GaAs benefits from a generous body of publications around the subject of dopant diffusion. For the most part, those studies have shown a concentration-dependence for Si diffusion in GaAs and a compensation effect at high dopant level, similar to results in InGaAs. However, the atomistic origin of those phenomena were still not completely known. Early hypothesis have relied on pure thermodynamics arguments and models built thereby relied on very coarse assumptions regarding mobile species. In this chapter, using NEB calculations, we provide the critical missing part of kinetic considerations by exploring the energy barriers associated with a number of vacancy-assisted and interstitial transitions. In light of those new data, we evaluate the limitations of current models. We also show for the fist time that interstitial silicon species can adopt split interstitial configurations which have been shown to be lead to highly mobile As species in GaAs. In the next chapter we will present results on our diffusion studies in InAs.

CHAPTER 4

INTRINSIC AND SILICON DIFFUSION IN InAs

4.1 Introduction

Among the III-V materials, InAs stands out because of its very high electron mobility which can be as much as three times higher than in InGaAs and GaAs [28, 96]. It also has a small direct band gap of 0.35 eV at room temperature and low carrier effective masses [97]. Together, these properties make InAs a promising candidate for incorporation into next-generation nanoelectronics [28]. In addition, InAs has successfully been made into nanowires [98, 99, 100] and it has been demonstrated to integrate well into novel field effect transistor (FET) device geometries [101, 102, 103, 38].

Due to the importance of surface states in nanowires, numerous previous studies of defects in InAs have focused on understanding the role of surface states in the performance of nanowire-based InAs devices [104, 105, 106]. Overall, these studies have shown that surface crystallographic orientation and passivation strategy of dangling bonds can have a non-negligible impact on the effective masses and mobilities of the carriers. Their results have provided valuable understanding of the effect of surface states on the electronic properties of InAs devices. However, in order to develop protocols intended to optimize fabrication processes, the community will also need a sound understanding of de-

This chapter was published in Langmuir [95]

fect energetics and diffusion in bulk InAs, which is not currently available. Several key unanswered questions remain, such as: What is the primary diffusion mechanism of Si, the most common dopant in InAs? And what are the relative stabilities of interstitial and substitutional defects? The answers to these questions will affect the efficacy of thermal annealing processes and alter the degree of Si dopant activation in the matrix (and hence conductivity); these answers cover two important aspects of semiconductor processing.

In this computational study, we focus primarily on providing answers to the first of those questions, namely to understand defect energetics within the bulk InAs crystal as they pertain to dopant diffusion. We considered the diffusion of both intrinsic defects and silicon (dopant) atoms within the InAs crystal lattice. We study silicon because it has historically been used as the premiere choice for n-doped InAs [37, 38, 107, 108, 109, 110]. For increased accuracy, our calculations are performed at the *ab initio* level using Density Functional Theory (DFT). We performed Nudged Elastic Band calculations for vacancy-assisted and interstitial transitions to determine the preferred diffusion routes of those defects and understand the kinetics of their diffusion.

4.2 Methods

We use Quantum Espresso [111] code for all the calculations performed in this study. While a hexagonal wurtzite structure has been found to arise in InAs nanowires, bulk InAs, like most III-V materials, typically adopts a zincblende structure [112]. Accordingly, we perform our calculations with a zincblende InAs supercell containing 64 atoms as shown in Fig. 4.1. Due to the large num-

ber of calculations involved in this study, simulating bigger system sizes was out of reach due to limited computational resources. Moreover, we have previously investigated the potential system size effects of using a 64-atom supercell in a closely related system, namely CuAu-I $In_{0.5}Ga_{0.5}As$, and found no significant changes in the energy barriers calculated as we increased the size of the cell [113]. Hence, we believe that system size effects are relatively small.

We found that a cut-off energy of 75 Ry and a 3x3x3 k-mesh achieved a targeted convergence of 0.005 eV or less in energy. The simulations were performed with norm-conserving Perdew-Zunger pseudo-potentials within the Local Density Approximation (LDA) level of theory. An LDA level of theory has been shown to be an appropriate choice for defect formation and migration energy calculations for III-V systems of this type [92, 93]. This has also been confirmed by some of our calculations (not reported here) that show negligible improvement when using GGA rather than LDA for Nudged Elastic Band (NEB) calculations in the closely related system of CuAu-I $In_{0.5}Ga_{0.5}As$.

After geometry optimization of the perfect crystal, we introduced defects by either removal of an appropriate atom to create a vacancy, or addition of an extra atom to create an interstitial in a pre-specified position (tetrahedral or hexagonal). The resulting structures are allowed to relax energetically, permitting atoms to move in all three dimensions. All geometry optimization are performed using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, as implemented in Quantum Espresso. We studied the following defects: In, As and Si hexagonal and tetrahedral interstitials (six case studies), In and As vacancies (V_{In} and V_{As}), as well as antisite and dopant substitutions In_{As}, As_{In}, Si_{In} and Si_{As}, where the subscript denotes the site on which the atom sits. For

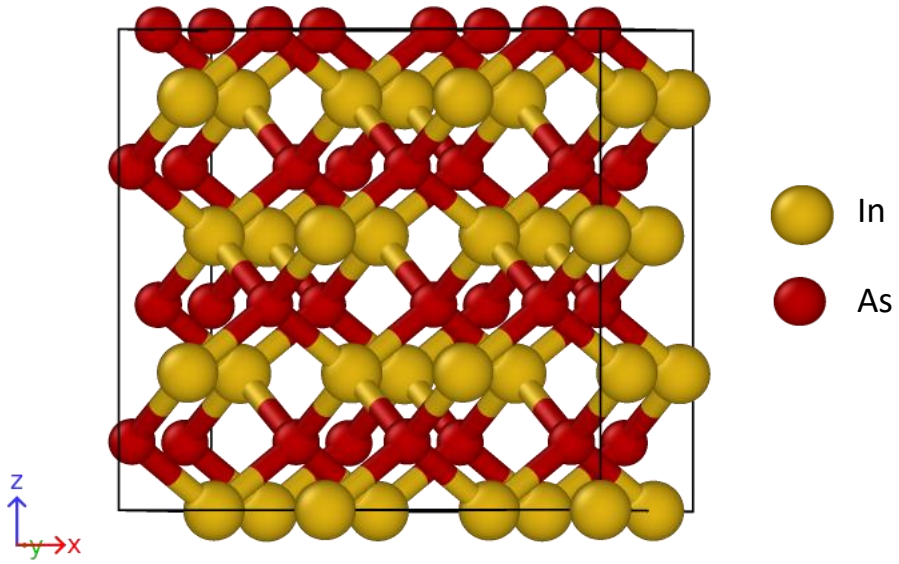


Figure 4.1: A representative 64-atom InAs supercell in the cubic Zincblende structure. Unless otherwise noted, indium atoms are shown in yellow and arsenic in red.

instance, In_{As} implies that indium atom occupies an arsenic site.

We used Nudged Elastic Band (NEB) calculations to probe the energetics of transitions involving the movement of the defects, described above, in the InAs bulk lattice. NEB is a well-known method used to find the Minimum Energy Path (MEP) between two chosen states of a system [57, 58, 59], *i.e.*, an initial and final state. Using this method, we are able to calculate the energy barriers required for our chosen defects to move within the InAs lattice. In this manner, we studied vacancy-assisted first-nearest neighbor transitions (between the In and As sublattices) and second-nearest neighbor transitions (within a given sublattice) as well as interstitial jumps.

To complement these diffusion pathway (migration energy) studies, we also performed limited formation energy calculations for the defects of interest in this work *in their neutral state*. A comprehensive study of defect formation energies would include investigations of these energies as a function of Fermi level and charge; but such a study is outside the scope of this paper. The formation energies are computed using the well-known approach described in Northrup *et al.* [114]. Within that framework, the formation energy, E^f, is given by Eq. 4.1,

$$E^f(D^q) = E(D^q) - E_{bulk} - \sum_i \Delta n_i \mu_i + q(E_V + E_F) + E_{corr} \tag{4.1}$$

$$E^f(D^q) = E(D^q) - E_{bulk} - \sum \Delta n_i \mu_i \tag{4.2}$$

where $E(D^q)$ is the total energy of the crystal with the defect, E_{bulk} is the

energy without the defect, Δn_i is the change in the number of atoms of species i, and μ_i is the chemical potential of species i. Since our calculations involve neutral defects, the correction term E_{corr} is not necessary. The term involving $q(E_v + E_F)$ accounts for the Fermi level and charge dependence of the formation energy; this can also be ignored here for neutral defects, leading to the simplified version given by Eq. 4.2. Since the exact value of the chemical potential cannot be determined, it is used as a parameter for the formation energy calculations. As such, the defect formation energies are given in the limiting conditions of As-rich and In-rich growth regimes. In the As-rich regime, the chemical potential of As is assumed to be its chemical potential in bulk As whereas in the In-rich (As-poor) regime, it corresponds to the chemical potential difference between InAs and bulk In (and vice-versa for the chemical potential of In). For an in-depth discussion of formation energy calculations, the reader is referred to the following papers. [92, 114, 115, 116]

4.3 Results and discussion

4.3.1 Vacancy-assisted diffusion

Energy barriers for all the vacancy-assisted transitions found in this study are summarized in Table 4.1. The energy barriers associated with In and As vacancy-assisted jumps in InAs are shown in Fig. 4.2 and 4.3, respectively. Our calculations show that if there is an In vacancy within one nearest neighbor distance, In will not stay on an As site. In other words, the $In_{As} - V_{In}$ antisite-vacancy defect complex is unstable and spontaneously relaxes back to

the $In_{In} - V_{As}$ single defect. This result is in contrast to the situation in which As occupies an indium site and a vacancy is nearby. In such a configuration, the antisite-vacancy defect complex ($As_{In} - V_{As}$) is stable. The transition that transforms this defect complex into a single In vacancy ($As_{As} - V_{In}$) has an energy barrier of 0.59 eV. Energy barriers for In and As moving within their sublattice are quite high, 1.67 eV and 2.00 eV, respectively. If As diffuses into an antisite, the barrier involved is 1.83 eV for $As_{In} \rightarrow As_{In}$, as reported in Table 4.1.

Turning from intrinsic defects to silicon dopants, we calculated NEB results for Si transitions involving vacancy-assisted jumps; see Fig. 4.4. The two transitions involving Si moving within either the In or the As layer have energy barriers of 2.52 and 2.18 eV, respectively. Second nearest neighbor jumps, such as those, are generally expected to have higher energy barriers because of the distance involved and the number of bonds that need to be broken and reformed in the process. However, it is important to note that these values for second nearest neighbor Si transitions are greater than their corresponding In and As counterparts. This suggests that dopant Si atoms will be the slowest vacancy-assisted diffuser in the InAs crystal.

Overall, our results for vacancy-assisted jumps show that As is the fastest and most "versatile" diffuser in Si-doped InAs. It can easily occupy In sites to form As_{In} defects even in the presence of an As vacancy within one nearest neighbor distance. In addition, energy barriers for As jumps can be as small as 0.59 eV for first nearest neighbor transitions. In the case of In, first nearest neighbor jumps from an In to an As site do not take place. Indium strongly prefers to stay on the In sublattice within which it can only move with relatively large energy barriers of 1.67 eV, as befits this considerably larger atom than

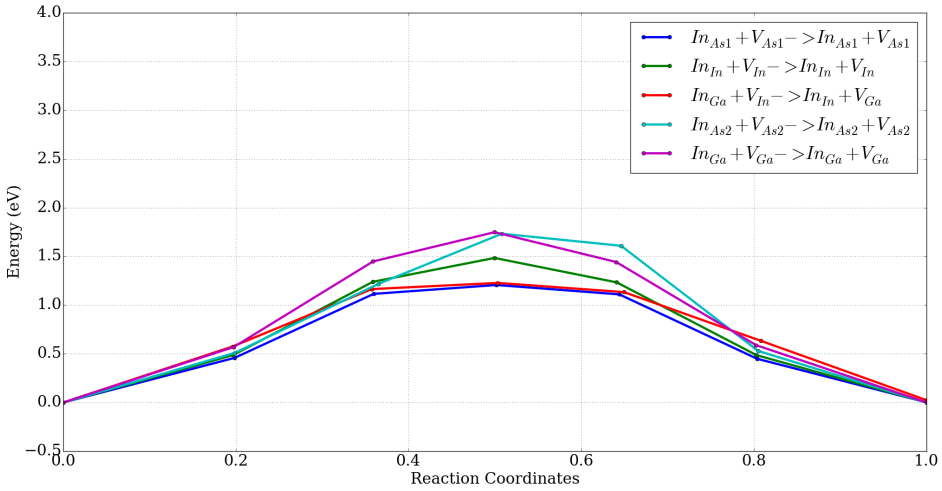


Figure 4.2: DFT-NEB results showing an energy barrier of 1.6 eV for In to jump to a nearby In vacancy in the InAs lattice. For improved clarity, the moving In atom is shown in green in the insets.

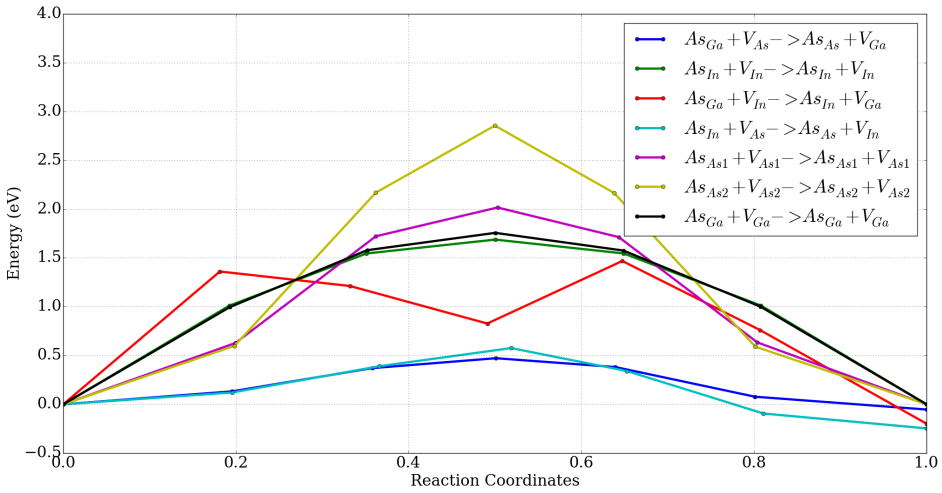


Figure 4.3: DFT-NEB results for vacancy-assisted As jumps in InAs. Arsenic can undergo transitions that takes it to both In and As vacancies alike. The insets show the initial, saddle point and final configurations of a first nearest-neighbor As jump (as represented by the green curve). The moving As atom is shown in blue for added clarity.

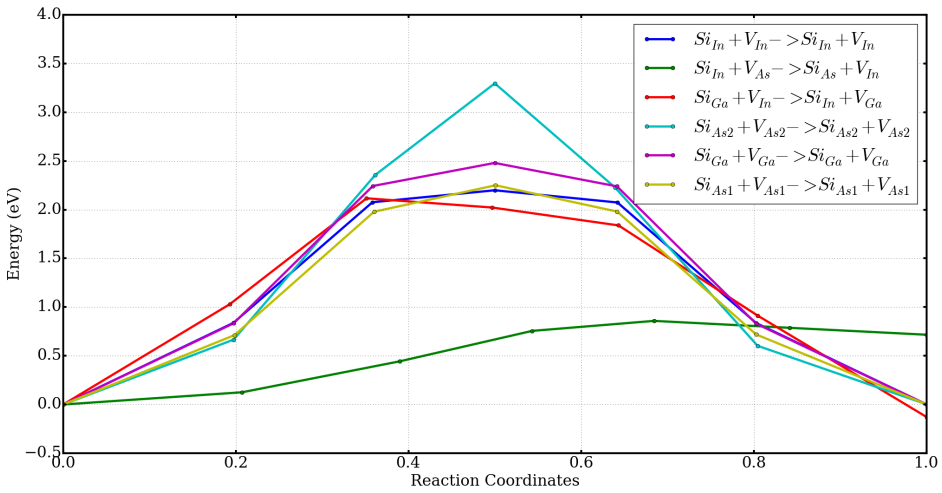


Figure 4.4: DFT-NEB results for vacancy-assisted Si jumps in InAs. The insets show the initial, saddle point and final configurations of Si (colored in blue for added clarity) moving between two In sites (green curve).

either As or Si. In contrast to As and In, vacancy-assisted Si jumps involve energy barriers as high as 2.52 eV, indicating that a vacancy-assisted mechanism is very unlikely to be the primary diffusion route for Si.

4.3.2 Interstitial diffusion

We also investigated the possibility of interstitial transitions, again using NEB calculations. Starting from atoms located in tetrahedral interstitial sites, we perform a geometry optimization, followed by an NEB calculation, as described in section 4.2. Our results, summarized in Table 4.2, show that interstitial In atoms preferentially occupy tetrahedral sites. An example of such configuration is given in Fig. 4.5. In contrast, As prefers to adopt split-interstitial configurations whereby a lattice site is shared with another As atom. One such configuration is shown in Fig. 4.6. the presence of split-interstitial As configurations have recently been reported in $In_{0.5}Ga_{0.5}As$ by Reveil *et al.* [113] and in GaAs by Wright *et al.* [93]. On the other hand, Si strongly prefers to occupy hexagonal sites, although tetrahedral configurations are also relatively feasible. An example hexagonal configuration for Si is shown in Fig. 4.7. The implications of Si occupying hexagonal versus tetrahedral or split-interstitial sites in InAs in terms of defect population and activation remains an open question. Subsequent calculations that incorporate the effects of charges are needed to answer such question.

Figures 4.11 and 4.12 show the NEB energy profiles for transitions involving interstitial In and As atoms, respectively. The energy barrier for In moving between tetrahedral configurations is 0.82 eV, which is considerably lower than

Atom Moving	Initial Lattice Site	Final Lattice Site	Energy Barrier (eV)
In	In	As	N/A
	As	In	No barrier
	In	In	1.67
	As	As	N/A
As	In	As	0.59
	As	In	0.59
	In	In	1.83
	As	As	2.00
Si	In	As	N/A
	As	In	No barrier
	In	In	2.52
	As	As	2.18

Table 4.1: Summary of energy barrier results from DFT NEB calculations for vacancy-assisted transitions in InAs. N/A refers to transitions that do not take place because it would involve a monotonic increase in energy of the system (with no barrier). Conversely, "No barrier" refers to a transition that goes downhill in the energy landscape.

the barriers obtained for vacancy-assisted jumps. For As, the barrier for jumps between split interstitial configurations is even lower, 0.41 eV. Such low energy barriers for As jumps are consistent with our results for vacancy-assisted diffusion showing that As is a very fast diffuser in InAs.

In Fig. 4.10, we plot the NEB energy profiles for Si jumps. The barrier to jump between hexagonal configurations is only 0.23 eV. This value is smaller than the energy barriers for all the other transitions computed thus far. For comparison purposes, the smallest energy barrier for vacancy-assisted Si jumps was 2.18 eV, almost 2 eV higher! This suggests that Si will almost certainly diffuse *via* an interstitial mechanism, rather than a vacancy-assisted one. Figure 4.10 also shows the energy profile for interstitial Si moving from an hexagonal to a tetrahedral configuration. As evident in the associated energy curve, the hexagonal configuration is strongly preferred to the tetrahedral one.

4.3.3 Defect formation energies

Figure 4.13 shows the formation energy as a function of chemical potential for the neutral point defects of interest in this work. We remind the reader that these defect formation energies are far from being comprehensive and are included here as a preview into the thermodynamics of those point defects in InAs. It is evident from this plot that As vacancies have low formation energy in InAs specially in the In-rich regime. Added to the relatively small energy barriers associated with As migration, this means that As defects will be very easy to form and migrate within the InAs crystal. The second group of defects with the lowest formation energy is comprised of $Si_{In}-V_{In}$ for As-rich growth conditions

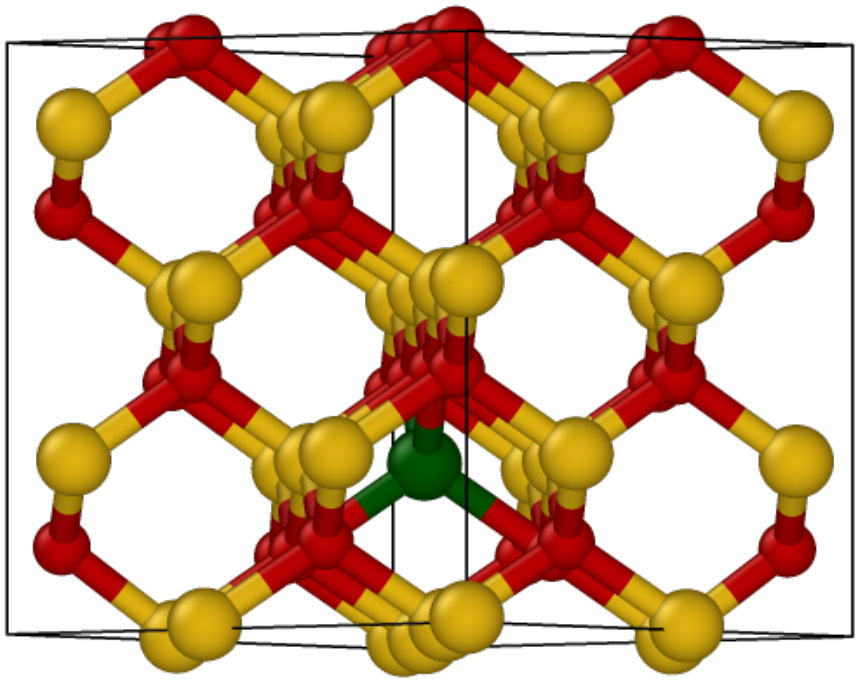


Figure 4.5: An indium atom (shown in green) in tetrahedral interstitial configuration in InAs.

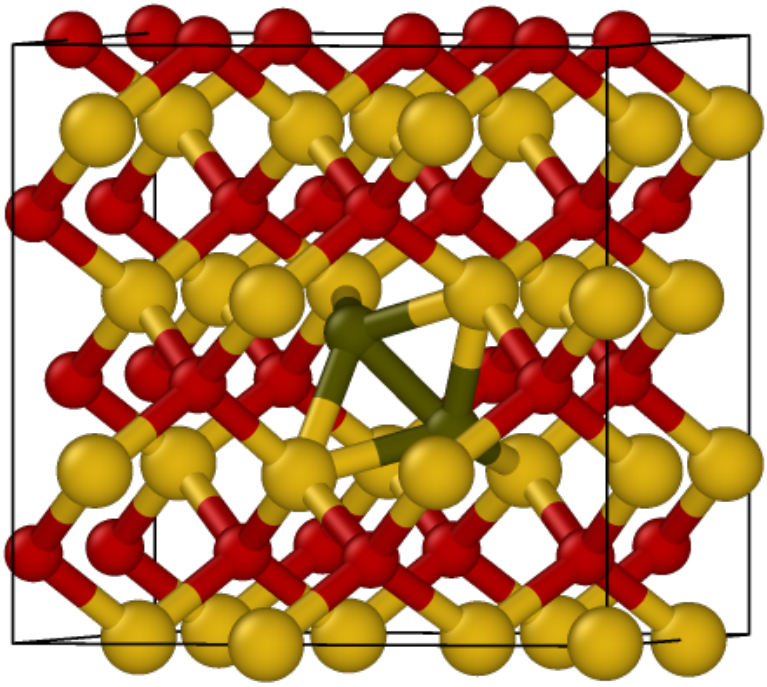


Figure 4.6: An arsenic split-interstitial configuration in InAs.

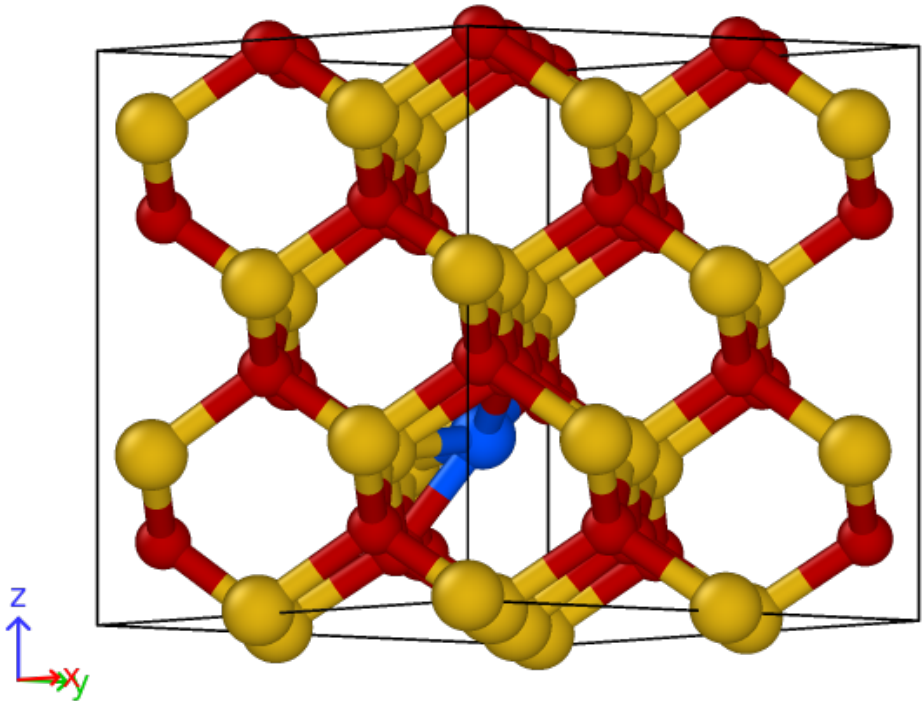


Figure 4.7: A silicon atom colored in blue in hexagonal interstitial position in InAs.

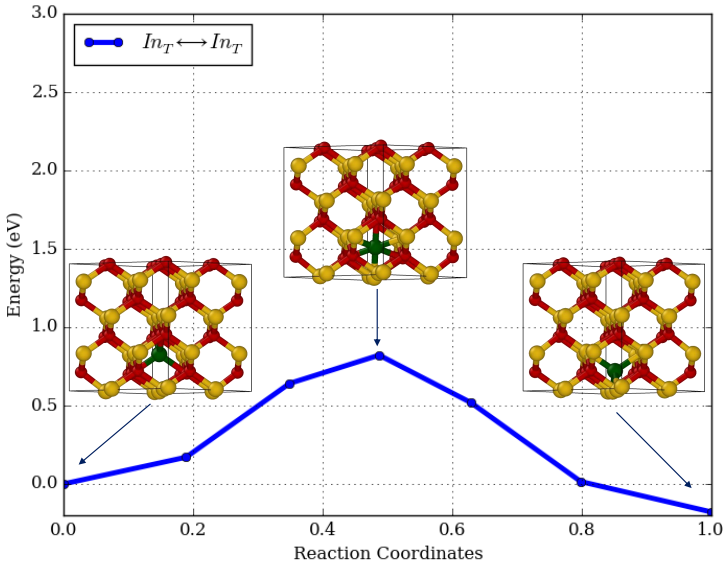


Figure 4.8: DFT-NEB results for In interstitial jumps in InAs showing a relatively low 0.8 eV barrier. The moving In atom is colored in green.The saddle point corresponds to an hexagonal configuration.

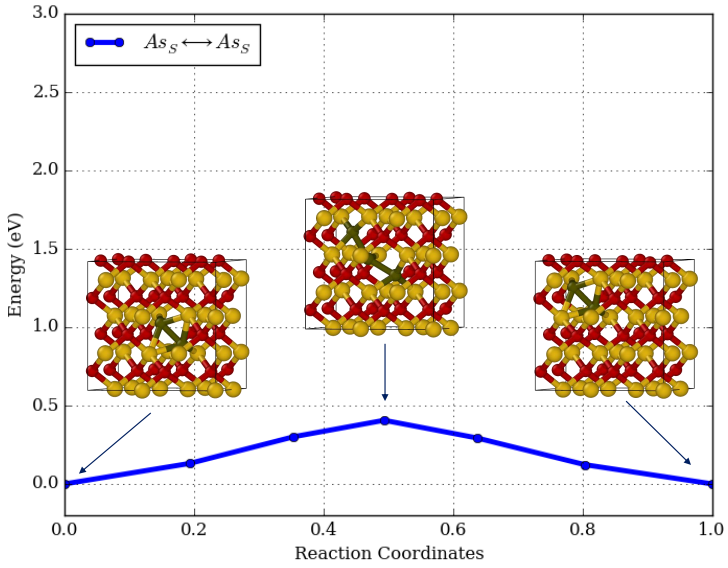


Figure 4.9: DFT-NEB results for As interstitial jumps in InAs, showing the ease of As diffusion in InAs (< 0.5 eV barrier) *via* an interstitial mechanism. The As atoms involved in the transition are shown in green for more clarity.

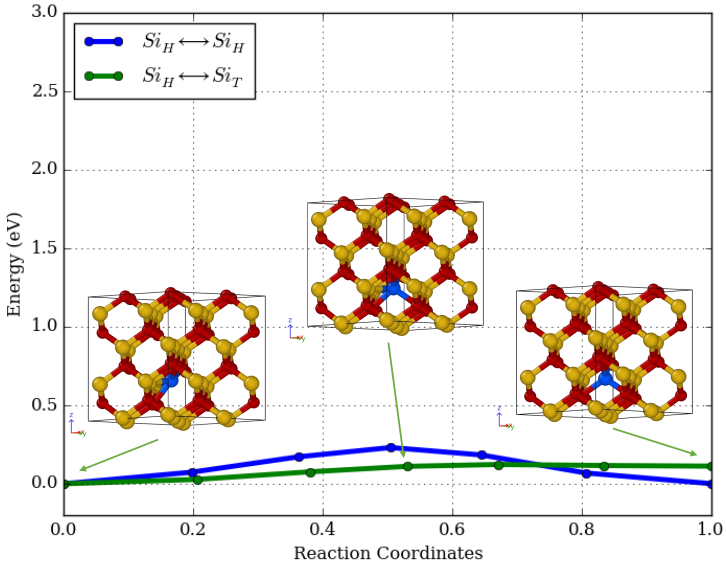


Figure 4.10: DFT-NEB results for Si interstitial jumps in InAs, showing the lowest energy barriers among the set of pathways we investigated, a < 0.25 eV barrier for hexagonal transport (blue curve) and showing that the hexagonal interstitial site is preferred to the tetrahedral one (green curve) by about 0.2 eV.The hexagonal to tetrahedral transition is illustrated by insets showing the initial, saddle point and final configurations with the moving Si atom colored in blue.

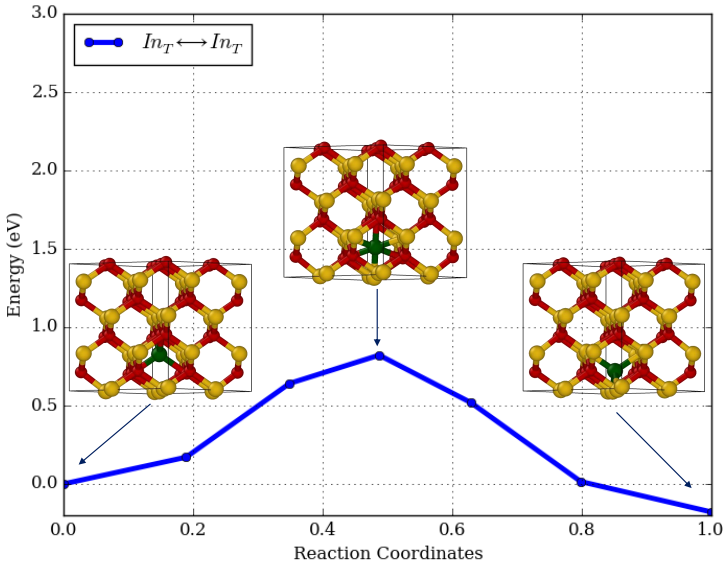


Figure 4.11: DFT-NEB results for In interstitial jumps in InAs showing a relatively low 0.8 eV barrier. The moving In atom is colored in green.The saddle point corresponds to an hexagonal configuration.

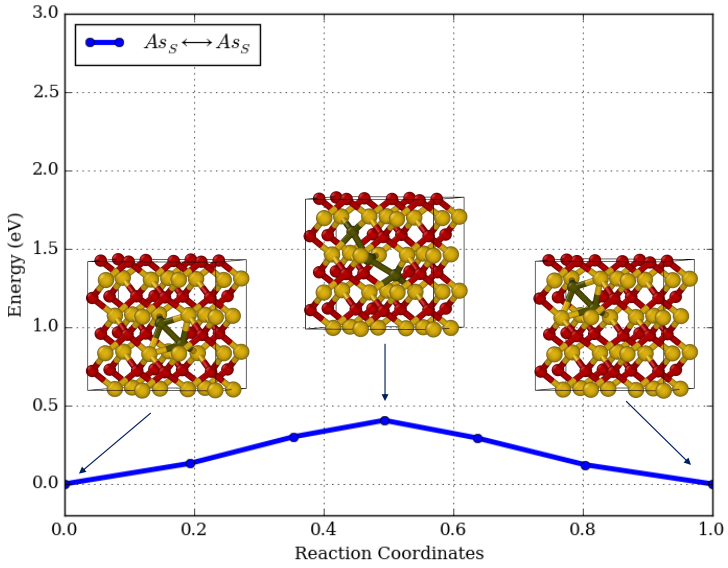


Figure 4.12: DFT-NEB results for As interstitial jumps in InAs, showing the ease of As diffusion in InAs (< 0.5 eV barrier) *via* an interstitial mechanism. The As atoms involved in the transition are shown in green for more clarity.

Atom Moving	Initial Interstitial Site	Final Interstitial Site	Energy Barrier (eV)
In	Tetrahedral	Tetrahedral	0.82
As	split-interstitial	split-interstitial	0.41
Si	Hexagonal	Hexagonal	0.23
	Hexagonal	Tetrahedral	0.11
	Tetrahedral	Hexagonal	0.001

Table 4.2: Summary of energy barrier results from DFT NEB calculations for interstitial transitions in InAs showing the generally lower barriers for interstitial jumps.

and $Si_{As} - V_{As}$ for In-rich growth conditions. On the other hand, as previously discussed, vacancy-assisted Si jumps have high energy barriers. In contrast, Si_H defects that have low energy barriers for jumps within the InAs lattice, but have the highest formation energy in this study. This means that, overall, unless the growth technique favors the formation of Si interstitials, Si is not expected to show significant diffusion in InAs. This would be deleterious for activation of Si dopants in InAs.

4.4 Conclusions

The results in this study are, for the most part, very similar to Reveil *et al.*'s findings for diffusion in $In_{0.5}Ga_{0.5}As$ (InGaAs) [113] and by Wright *et al.* for GaAs [93]. Taken together, these results provide a more comprehensive overview of diffusion across the more common III/V alloys. In both InAs and InGaAs, As is found to be a fast diffuser and to preferentially adopt split-interstitial configurations in which two As atoms share a lattice site and are able to move in the lattice with very low energy barriers. Not surprisingly, given its large size, indium is found to be a very slow diffuser, with energy barriers for vacancy-assisted jumps well above 1.5 eV. The diffusion of Si dopant atoms *via* interstitial diffusion is found to be more favorable than vacancy-assisted diffusion in both III-V materials.

However, there is one key difference between our results for $In_{0.5}Ga_{0.5}As$ versus InAs that stands out: Interstitial Si atoms prefer to adopt split-interstitial configurations in $In_{0.5}Ga_{0.5}As$, but do not in InAs. Instead, they prefer to occupy hexagonal sites. This suggests that the difference in size and/or electronegativ-

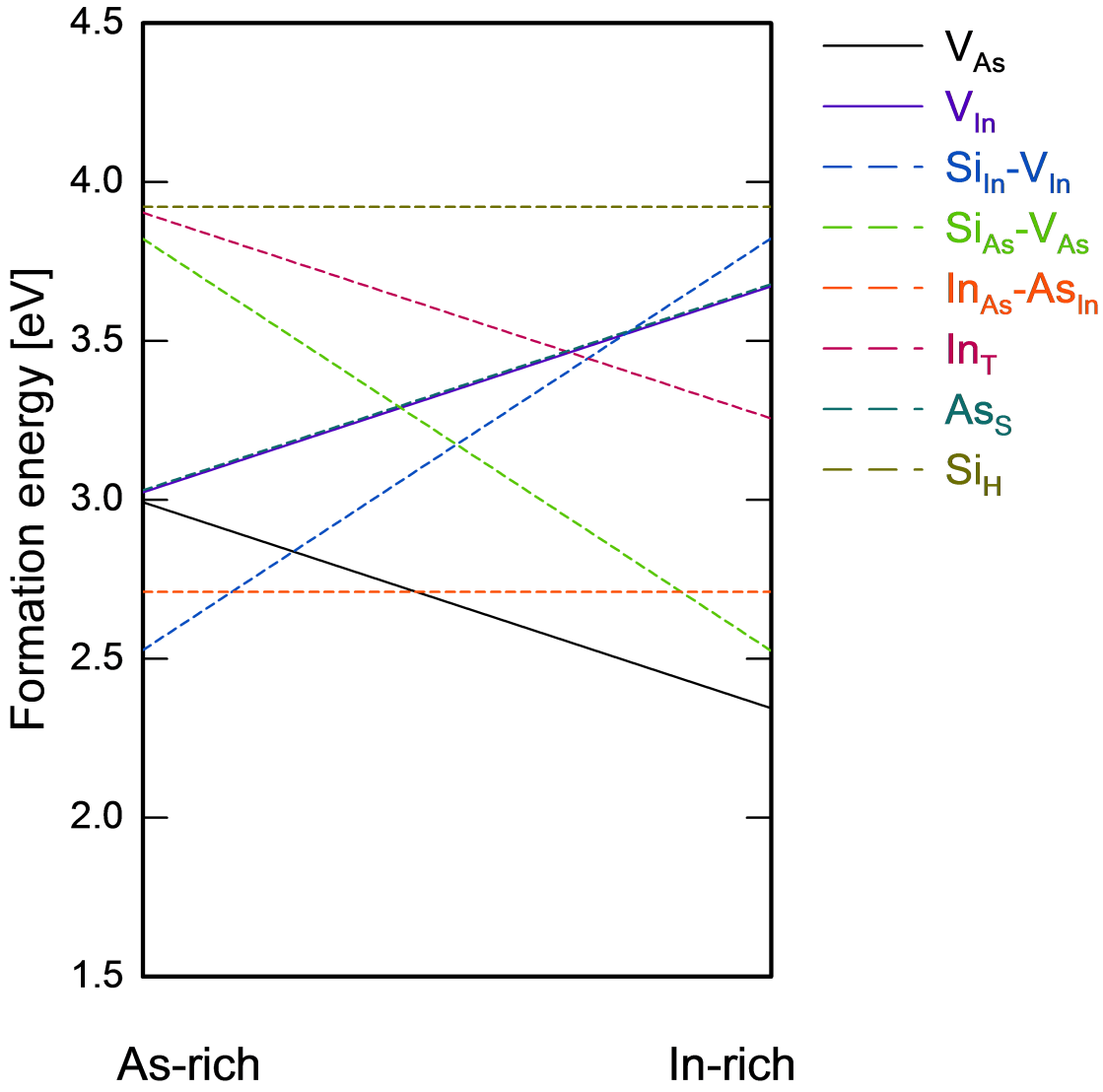


Figure 4.13: Formation energy of point defects in InAs with respect to growth conditions.

ity between In and Ga might influence the thermodynamic and diffusive behavior of group IV dopants like Si. The exact nature of such an influence beyond what type of site the Si dopant preferentially occupies, is still unknown and should be further investigated in those three related III-V materials (*i.e.* GaAs, InGaAs, and InAs).

CHAPTER 5

INTRINSIC AND SILICON DIFFUSION IN InGaAs

5.1 Introduction

As previously mentioned, silicon is a group IV atom and is, therefore, an amphoteric dopant in III-Vs. For n-type doping, Si must substitute for In or Ga on the cation sublattice. The substitution of Si for As on the anion sublattice has been considered as a potential compensation mechanism responsible for the observed activation limit [54]. However, recent papers have proposed an alternative explanation for both the activation limit [92] and the concentration-dependent diffusivity of Si in InGaAs [54, 117, 118]. Partly based on results for GaAs and data in InGaAs, these papers found that the formation energy of cation (group III) vacancies (V_{III}) decreases as the Fermi level rises with increasing Si concentration. At high Si concentration, electrically inactive $Si - Vacancy$ complexes are thought to form, which then lead to both enhanced diffusivity and a limit in activation.

In this paper, we use accurate Density Functional Theory calculations to understand the previously unknown kinetics of Si dopant behavior in InGaAs and to test the presumed importance of a vacancy-assisted mechanism for Si diffusion. We calculate energy barriers for more than 40 different transitions of silicon dopant atoms and intrinsic defects within the InGaAs crystal lattice. We

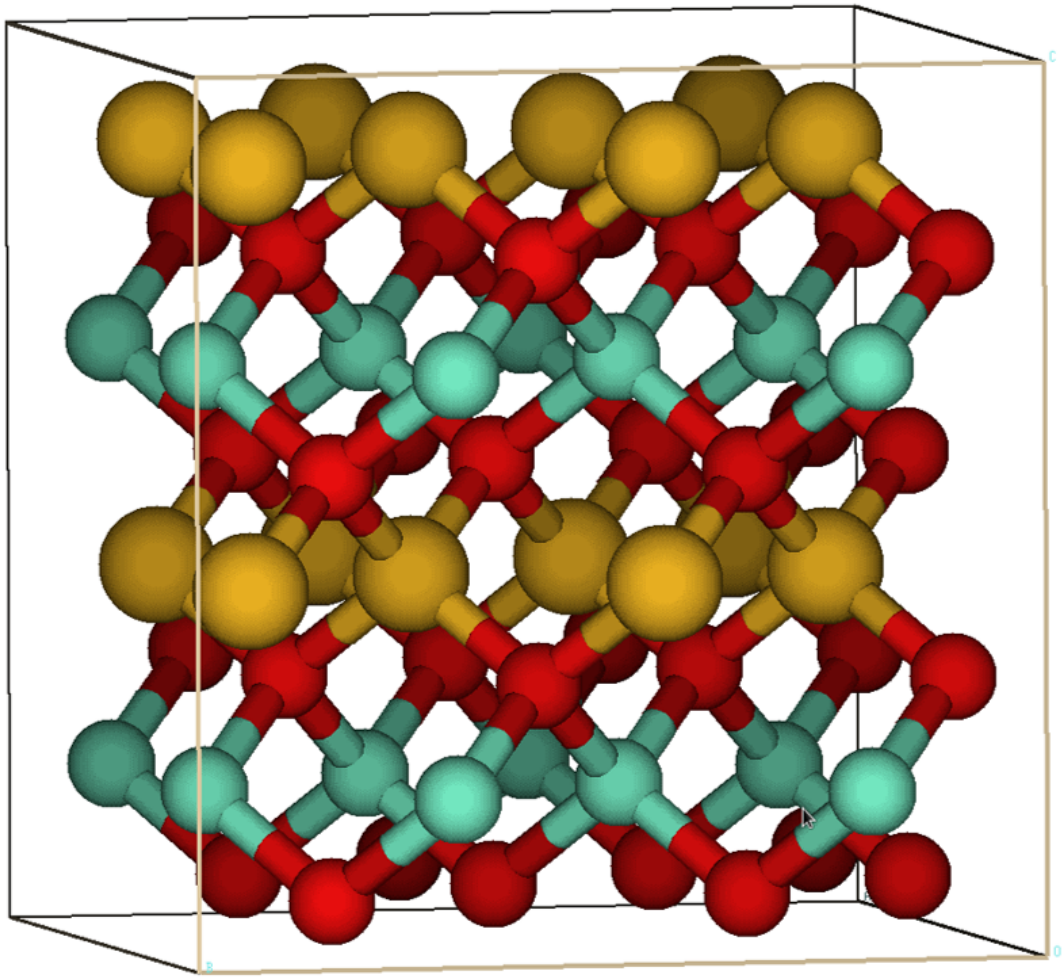
This chapter was published in Acta Materialia [113]

explore all possible vacancy-assisted transitions for In, Ga, As and Si in order to test the validity of a vacancy-assisted mechanism being the primary diffusional mechanism. We also study interstitial defects and dopants to complete our study of preferred mechanistic pathways. Comparing all these calculated diffusional rates allows us to provide new insights into the most favorable diffusional mechanisms at play. These new data represent the most extensive molecular-level theoretical study of Si diffusion in InGaAs to date.

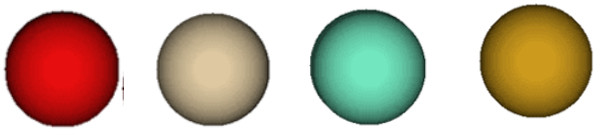
5.2 Methods

5.2.1 Crystal structure and ordering

Although it is known that InGaAs adopts a diamond cubic structure, the cation sublattice in InGaAs can adopt random or ordered arrangements, depending on growth conditions [119]. While random ordering might be commonly expected, evidence of CuPt-like ordering have been reported in many III-V alloys during low-temperature vapor phase epitaxy growth (VPE) [120, 121, 122]. CuAuI ordering has also been observed in III-V alloys, including InGaAs, during molecular beam epitaxy (MBE) [123, 124, 125, 126]. In this work, due to the inherent complexity and the prohibitively large system size necessary to study a random crystal, we focus on the CuAuI-ordered phase of InGaAs, shown in Fig. 5.1. In this ordering, the cation sublattice is comprised of alternating layers of In or Ga atoms stacked along the z-direction. However, as our results will show, we will have something to say about the likelihood of randomizing species swops (*e.g.*, swopping In for Ga).



(a)



(b)

Figure 5.1: (a) Schematic of a 64-Atom CuAuI-ordered $In_{0.5}Ga_{0.5}As$ supercell, showing alternating layers of atoms. The CuAuI ordering of InGaAs has been observed for specific growth conditions, *e.g.*, MBE growth. (b) Color key used throughout this text. From left to right: As, Si, Ga and In (Images not on scale)

5.2.2 Mechanisms studied

We have studied both vacancy- and interstitial- mediated transitions. Table 5.1 lists the vacancy-assisted transitions studied in this work. We performed calculations for both first- and second- nearest neighbor vacancy-assisted jumps for Si as well as vacancy-assisted moves of In, Ga and As atoms. We considered all three types of vacancies (missing an In, Ga or As atom) and explored all possible combinations of mobile atom/vacancy pairs in order to find the lowest energy transitions among this class of moves. For all calculations involving vacancies, the starting configurations were obtained by choosing one atom to remove from an initially perfect 64-atom supercell in order to form a vacancy. The atomic positions inside the supercell with the vacancy were then relaxed to a lower energy configuration.

We also studied the barriers to *interstitial* transitions for all four atoms: In, Ga, As and Si. Although we initially assumed a tetrahedral interstitial configurations, each supercell was allowed to relax and adopt a more stable configuration which, in some cases, was different from the tetrahedral starting point. The starting configurations were obtained by adding one extra atom to a 64-atom supercell in a specific tetrahedral position, followed by geometry optimization.

5.2.3 Simulation details

For each transition studied, *e.g.*, the diffusion of a vacancy from one lattice site to another, the related energy barrier was calculated using the Nudged Elastic Band (NEB) method. NEB is a widely used technique for finding the Minimum Energy Path (MEP) between two known configurations of a molecular system

Vacancy-Assisted Transitions			
	Atom Moving	Initial Lattice Site	Final Lattice Site
First Nearest Neighbor Jumps	In, Ga, As, Si	In As Ga As	As In As Ga
Second Nearest Neighbor Jumps	In, Ga, As, Si	In Ga In Ga As	In Ga Ga In As

Table 5.1: List of vacancy-assisted transitions studied in this work.

[58, 127, 128, 59]. The requirement that both the initial and final configurations are known in advance is not a complicating issue for on-lattice systems such as InGaAs.

The NEB module in Quantum ESPRESSO [111] was used for all NEB calculations. A 64-atom supercell of CuAuI-ordered InGaAs was used with a norm-conserving pseudopotential and Local Density Approximation (LDA) functional in the Perdew-Zunger (PZ) parameterization. The LDA level of theory has been previously shown to perform well for this class of system [92, 93]. Additional calculations that we performed at the GGA level confirmed such findings. A kinetic energy cut-off of 680 eV was used along with a $3 \times 3 \times 3$ k-mesh. These choices of pseudopotential, cut-off energy and k-mesh were made following appropriate tests to assess suitable convergence limits had been achieved. Due to the computational cost of *ab initio* NEB calculations and the large number of simulations performed in this study, the use of a larger supercell containing, say, 216 atoms was out of reach. Moreover, preliminary calculations that we performed with larger supercells have indicated that no significant accuracy gain is achieved by using 216- versus 64-atom supercells.

Before each NEB calculation, the initial and final configurations were relaxed to minimize the energy. Geometry optimizations were performed using the Broyden - Fletcher Goldfarb Shanno (BFGS) algorithm with convergence thresholds of 10^{-4} and 10^{-3} for the total energy and force, respectively, in atomic Rydberg units. For higher accuracy, the self-consistent convergence threshold was set to a value of 10^{-8} Ry which is two orders of magnitude lower than the commonly used value of 10^{-6} Ry. Relaxing the initial and final configurations before performing NEB calculations allowed us to identify unstable configu-

rations as well as situations when both configurations relax to the same final structure. This information provided valuable insights for subsequent calculations.

When NEB calculations led to more than one saddle point, the potentially stable intermediate configurations were isolated and relaxed using an energy minimization technique. The original NEB calculation was then broken down into multiple NEB runs, ensuring that there is only one saddle point per NEB calculation. More details on this process are given in Section 1 of the Supplemental Information. This meticulous approach allowed us to discover unsuspected stable configurations, as well as the discovery of low-energy transitions which might otherwise have been overlooked.

5.3 Results and discussion

In section 5.3.1, we present an overview of our results for vacancy-assisted diffusion and discuss their implications. In section 3.2, we show our results and discussion for interstitial diffusion in InGaAs.

5.3.1 Vacancy-assisted diffusion

As one of the most basic diffusion mechanisms in crystals, vacancy-assisted jumps allow atoms to leave their lattice site and migrate to a nearby vacant lattice site. In a ternary alloy such as CuAuI-ordered $In_{0.5}Ga_{0.5}As$, there are about 24 such transitions (not counting defect-defect interactions). Those transitions involve the movement of defects such as substitutional Si (Si_X), vacancies (V_X)

and anti-sites (X_Y) - where $X, Y = In, Ga, As$ and $X \neq Y$. We first consider first nearest neighbor jumps (section 5.3.1), followed by second nearest neighbor transitions (section 5.3.1). Finally, we provide an in-depth discussion in section 5.3.1 of the meaning and implications of the data.

First nearest neighbor jumps

First nearest neighbor jumps in InGaAs always take place between cation and anion sites and will, therefore, involve the formation of anti-sites in the case of intrinsic jumps. Depending on the stability of the defects created, we expect some of those transitions to be energetically unreachable or to have some preference for one of the sub-lattices. Table 5.2 shows a summary of all the results for first nearest-neighbor jumps. In-depth analysis of those transitions is included in section 2.1 of the Supplementary Information.

Indium and gallium do not move by first nearest neighbor jumps as those jumps would create In_{As} and Ga_{As} anti-sites which are unstable when a Ga or In vacancy is present as a first nearest neighbor. "Unstable" here means they will spontaneously leave that configuration with no energy barrier. However, As_{Ga} and As_{In} anti-sites *are* stable (meaning there is an energy barrier preventing them from spontaneously leaving that configuration) and first nearest neighbor As jumps take place with energy barriers between 0.47 eV to 0.82 eV. As expected, since Si is expected to lead to n-type doping, Si_{Ga} and Si_{In} defects are more favorable compared to Si_{As}, even when a vacancy exists within the first nearest neighbor "shell." The strong preference of Si dopants for cationic sites in InGaAs has been previously reported [115]. It takes only 0.14 eV for a Si atom to move from an As site to an In site, and there is no energy barrier at all to a Ga

vacancy. The reverse transition costs 0.86 eV in the case of In, and does not take place in the case of Ga.

Second nearest neighbor jumps

Based on the symmetry of the CuAuI ordering of InGaAs, second nearest neighbor jumps happen only between cationic or anionic sites. Except for jumps between In and Ga sites, the NEB results for these transitions are therefore perfectly symmetric. Energy barriers for second nearest neighbor jumps are expected to be higher than those for first nearest jumps. This is because second nearest neighbor jumps take place over longer distances and involve more bond breaking/forming compared to first nearest neighbor jumps. As shown in Table 5.3, all second nearest neighbor jumps in our calculations have energy barriers greater than 1 eV. In the Supplemental Information, we provide detailed information on Ga and In transitions (section S 2.2.1), As jumps (section S 2.2.2) and Si transitions (section S 2.2.3).

Discussion on vacancy-assisted diffusion

Taking all these preceding NEB results as a whole suggests that the species most likely to diffuse by vacancy-assisted transitions in InGaAs is As. This is due, in part, to the fact that anti-sites with As on cation sites are stable and can potentially play an important role during diffusion. As discussed above, In and Ga do not undergo first nearest neighbor jumps as these jumps would lead them to an unstable configuration involving an anti-site on As. In addition, since the energy barriers for In or Ga second nearest-neighbor jumps are on the order

Atom Moving	Initial Lattice Site	Final Lattice Site	Energy Barrier (eV)
In	In	As	N/A
	As	In	No barrier
	Ga	As	N/A
	As	Ga	No barrier
Ga	In	As	N/A
	As	In	No barrier
	Ga	As	N/A
	As	Ga	No barrier
As	In	As	0.57
	As	In	0.82
	Ga	As	0.47
	As	Ga	0.52
Si	In	As	0.86
	As	In	0.14
	Ga	As	N/A
	As	Ga	No barrier

Table 5.2: Summary of energy barrier results from DFT NEB calculations for first-nearest neighbor vacancy-assisted transitions in CuAuI-ordered InGaAs. "N/A" indicates a transition that does not take place as it would lead to a high energy configuration without involving an energy barrier. The reverse case is labeled "No barrier" and refers to a transition that takes place spontaneously from a high energy configuration to a low energy configuration with no energy barrier.

Atom Moving	Initial Lattice Site	Final Lattice Site	Energy Barrier (eV)
In	In	In	1.48
	Ga	Ga	1.75
	In	Ga	1.20
	Ga	In	1.23
	As	As	N/A
Ga	In	In	1.83
	Ga	Ga	2.07
	In	Ga	1.61
	Ga	In	1.57
	As	As	N/A
As	In	In	1.69
	Ga	Ga	1.76
	In	Ga	1.67
	Ga	In	1.47
	As	As^{Ga}	2.02
	As	As^{In}	2.86
Si	In	In	2.20
	Ga	Ga	2.48
	In	Ga	2.24
	Ga	In	2.12
	As	As^{Ga}	2.25
	As	As^{In}	2.30

Table 5.3: Summary of energy barriers for second-nearest neighbor vacancy-assisted transitions in CuAuI-ordered InGaAs. The superscript denotes the common nearest neighbor between the two As lattice sites in the case of As and Si jumps.

of 2.0 eV, whenever there is a Ga or In vacancy, an As atom will preferentially make a first-nearest neighbor jump to occupy that vacancy, incurring energy barriers as low as 0.52 eV for Ga and 0.82 eV for In. Those As jumps will essentially transfer cation vacancies to the As sub-lattice, making them even less available for second nearest neighbor jumps on the cation sub-lattice.

Based on this observation, one potential route for Si diffusion (in case there is a Ga or In vacancy site nearby) could be first nearest neighbor jumps of Si_{III} to As vacancies followed by subsequent jumps on the As sublattice until another Ga or In vacancy is found. To partly test this idea, we allow a configuration with a Si_{III}, a As_{III}, and a V_{As} to relax in order to understand how the vacancy site could be transferred from the Ga/In sublattice to the As sublattice where it would be available for first nearest neighbor Si jump. However, as the results presented in Fig. 5.2 show, when a Si atom is present as a second nearest neighbor of a Ga or In vacancy, first nearest neighbor As jumps to that vacancy to form a $Si_{III} - V_{As}$ complex become impossible. This is because the configuration involving an As_{III} anti-site, a Si_{III} substitutional dopant, and an As vacancy (essentially the complex $As_{III} - V_{As} - Si_{III}$) is highly unstable and will relax back to a $Si_{III} - V_{III}$ complex pair.

These additional results suggest that the $Si_{III} - V_{III}$ pair is energetically favorable and that a mechanism involving Si jumps to As sites is improbable. It also provides evidence for a potential deactivation mechanism of Si in InGaAs whereby oppositely charged V_{III} defects remain close to Si_{III} defects leading to the Si activation limit consistently reported in the literature [54, 45, 46].

The fact that As can easily occupy cationic as well as anionic sites also provides an additional explanation as to why As loss might be so common during

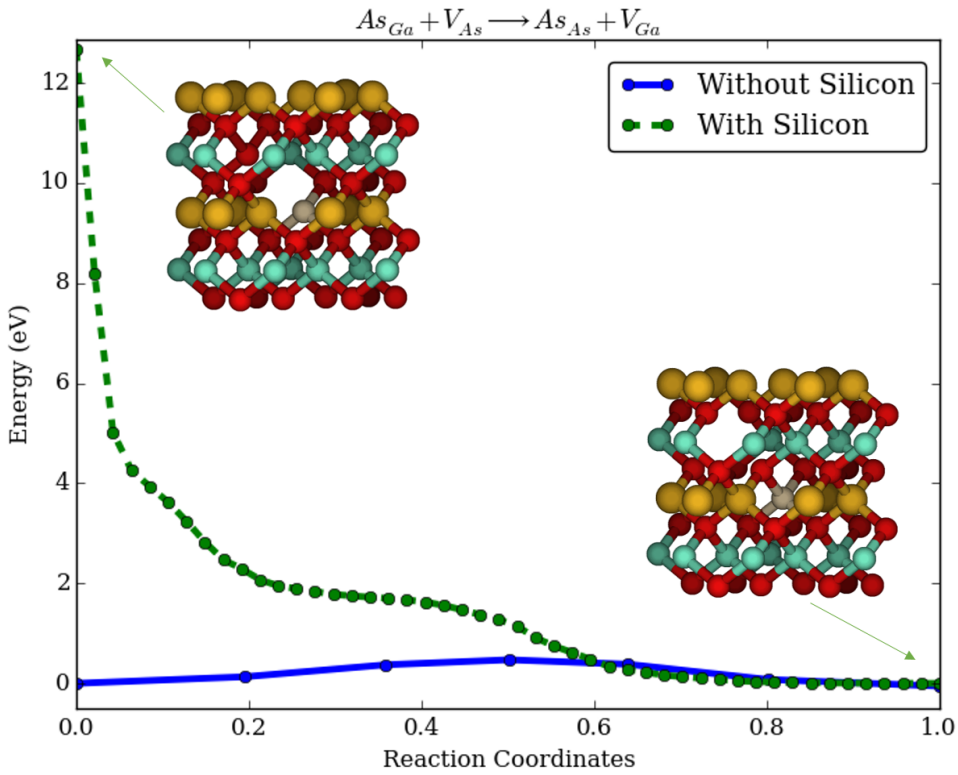


Figure 5.2: Energy profiles for As moving from a Ga site to an As vacancy with (green dotted line) and without (blue solid line) a Si atom as second nearest neighbor to a Ga site. The presence of Si renders the As_{Ga} defect unstable. The curve with silicon corresponds to a geometry optimization of the initial configuration which is shown to relax to the final configuration with no barrier. The curve without silicon corresponds to an actual NEB calculation as this transition involve an energy barrier.

annealing of InGaAs, as reported by several authors [46, 129, 130]. Slow transitions involving As atoms are indeed the most probable mechanism of vacancy-assisted intrinsic diffusion in CuAuI-ordered InGaAs. Si diffusion by vacancy-assisted jumps is possible but extremely slow, due to a combination of a high energy barrier for such jumps, a low energy barrier of As jumps to cation vacancies, and the stability of the $Si_{III} - V_{III}$ pair.

These results also highlight fundamental differences in the way Ga/In vacancies behave in InGaAs compared to As vacancies. The former seem to induce strain in the host lattice conducive to attracting Si dopants in their vicinity. This is evidenced by the stability of the $Si_{III} - V_{III}$ pair, as discussed above. From that perspective, Ga and In vacancies appear to behave as sinks for other point defects. This is probably attributable to the difference in atomic volume between Ga/In and As and the symmetry of the CuAu-I crystal form. In contrast, As vacancies do not seem to exhibit similar behavior. Further studies into the strain effects of point defects in InGaAs could help clarify the extent of such effects and how they influence diffusion and composition in their vicinity.

5.3.2 Interstitial diffusion

Interstitial sites in CuAuI-ordered InGaAs

We have explored a number of possible interstitial sites in InGaAs. Three different tetrahedral sites, which we have named T_1^a, T_1^b and T_2, are accessible based on the types of first and second nearest neighbors. As shown in Fig. 5.3, T_1^a corresponds to a tetrahedral site on a Ga layer with four As atoms as first nearest neighbors, four Ga and two In as second nearest neighbors. Similarly, T_1^b refers

to a tetrahedral site on an In layer with four As atoms as first nearest neighbors, four In and two Ga as second nearest neighbors. Finally, T_2 refers to a tetrahedral interstitial on an As layer with two In and two Ga as first nearest neighbors and six As atoms as second nearest neighbors.

Our calculations have revealed that In and Ga atoms can occupy tetrahedral interstitial sites, but pure interstitial jumps for As atoms do not take place. Instead, interstitial As jumps involve some type of split interstitial configuration. We have identified three such configurations, as shown in Fig. 5.4, which we have named S_1, S_2 and S_3. The first one, S_1, involves two As interstitials sitting on the same layer and sharing a common As lattice site. S_2 is very similar to S_1, except that the two As atoms occupy different layers. The last one, S_3, involves a slightly displaced In interstitial with As practically sitting on an hexagonal site (See Fig .5.4c). These configurations are completely different to the In and Ga tetrahedral interstitials.

In a similar fashion to As, interstitial Si atoms tend to adopt split interstitial configurations in which one As lattice site is shared between the Si atom and a displaced As atom. The possible configurations of Si split interstitials are shown in Fig. 5.5 and are similar to the ones presented in Fig. 5.4 for As. The S_2 configuration is energetically unfavorable for Si split-interstitials.

Energy barriers for interstitial jumps

Table 5.4 summarizes results for In, Ga, As and Si jumps between these interstitial sites. The energy barriers for interstitial jumps are, in general, lower than the vacancy-assisted ones. As mentioned above, In and Ga can move between

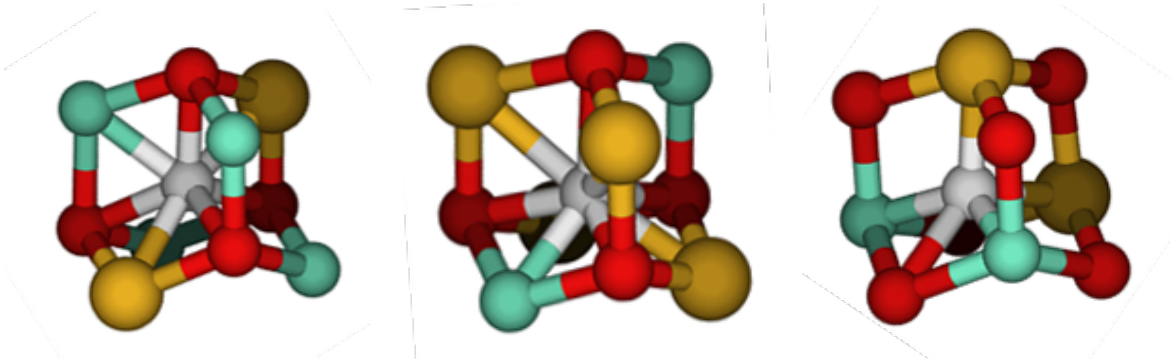


Figure 5.3: The three possible interstitial sites in CuAuI-ordered $In0.5Ga0.5As$. Left: T_{1a} with four As as first nearest neighbors (NN), four Ga and two In as second NN. Center: T_{1b} with four As as first NN, four In and two Ga as second NN. Right: T_2 with two In and two Ga as first NN, six As as second NN.

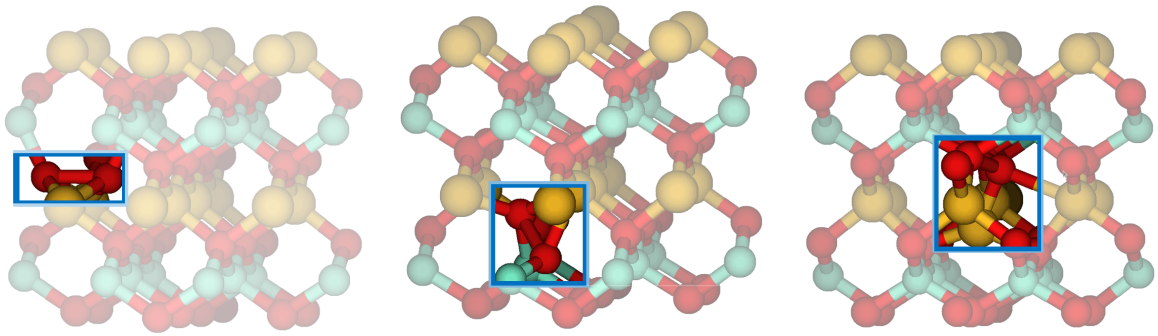


Figure 5.4: Examples of As split interstitial configurations. Left: S_1 configuration with two As atoms sharing an As lattice site. Center: S_2 configuration with two As atoms sharing a lattice site with a slight offset in the z direction. Right: S_3 configuration with an As sitting in an hexagonal position leading to slightly displaced In atom.

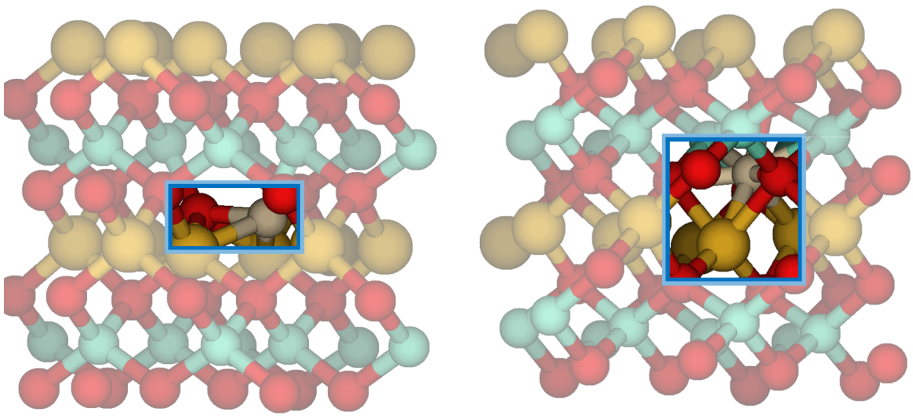


Figure 5.5: Examples of Si split interstitial configurations. Left: S1 configuration with a Si and an As atom sharing an As lattice site. Right: S3 configuration with a Si sitting in an hexagonal position leading to slightly displaced In atoms.

tetrahedral interstitial sites, whereas As and Si move between split interstitial configurations. The lowest energy for tetrahedral jumps is 0.63 eV, whereas the lowest energy for split interstitial transitions is just 0.13 eV. A more detailed description of those transitions is included in section 2.3 of the Supplemental Information.

Discussion regarding interstitial diffusion

Our results show that, while In and Ga tend to adopt tetrahedral interstitial positions, As and Si interstitials are found to prefer split interstitial configurations in which an As lattice site is shared. Since Si atoms in split interstitial configurations do not contribute to charge carrier concentration, our results suggest that split interstitials are a third possible origin for the observed activation limit of Si in InGaAs (the previous two options involved Si_{As} substitutional atoms and $Si_{III} + V_{III}$ complexes which would also contribute to dopant deactivation). The relative distribution of the population among these three species, with the propensity to limit dopant activation in InGaAs, will depend on the method used for dopant implantation as well as the Fermi level and temperature used during annealing.

Although we do not claim to have explored all possible split interstitial transitions that might occur, these results, involving favorable energy barriers as low as 0.13 eV for some As jumps, implicate As as the fastest diffusing species in InGaAs. In all our calculations, including those for vacancy-assisted diffusion, As has smaller energy barriers overall than all the other atoms. Therefore, any robust diffusion model for Si in InGaAs needs to account for As diffusion which can influence the overall intrinsic defect population.

Atom Moving	Initial Interstitial Site	Final Interstitial Site	Energy Barrier (eV)
In	T_2	T_1^a	1.25
	T_1^a	T_2	0.63
	T_2	T_1^b	0.9
	T_1^b	T_2	0.86
Ga	T_2	T_1^a	1.07
	T_1^a	T_2	0.77
	T_2	T_1^b	0.87
	T_1^b	T_2	0.82
As	S_1	S_1	1.44
	S_2	S_2	0.13
	S_3	S_3	0.13
	S_2	S_3	0.55
	S_3	S_2	0.16
Si	S_1	S_1	0.33
	S_1	S_1	0.42
	S_3	S_3	0.35
	S_1	S_3	0.15
	S_3	S_1	0.4

Table 5.4: Summary of energy barriers for interstitial transitions in CuAuI InGaAs. Two different paths for S_1 to S_1 Si transitions were explored.

As a complement to these migration energy studies, it is also helpful to understand the thermodynamic tendency to form Si split interstitial species. Thus, we have calculated their formation energy in the case of grown-in Si-doped CuAuI-ordered $In_{0.5}Ga_{0.5}As$. The procedure used for these calculations is the same as that given in Wang *et al.* [92], where the computational procedure was given in detail. The formation energies of Si split interstitials S_1 and S_3 vary between 2.5 and 3.7 eV, and remain stable at high Fermi level, see Fig. 5.6. In line with other ordered-InGaAs interstitial defects [92], S_1 and S_3 have formation energies that are independent of growth conditions (in terms of chemical potentials for In, Ga, and As). The most stable charge state of those split interstitials is neutral or +2 for Fermi levels within the band gap. Since the formation energy of Si substitutional atoms increases with Fermi level [92], while that of split interstitials remain constant, the relative population of Si split interstitials compared to Si dopants on substitutional sites is expected to increase at high dopant concentration, in line with experimental observations.

The fact that Si split interstitials are stable species and their diffusion has such low energy barriers is an unprecedented result in InGaAs. This discovery opens up a new possible model for Si diffusion in InGaAs, in which interstitial species are the ones diffusing at high concentration while substitutional atoms are immobile. It is known that ion implantation, for example, often leads to the presence of excess interstitial defects at high dopant concentration. It is not hard to imagine that, in such cases, interstitial atoms will be present near the surface where the dopant concentration is high and those atoms will be able to diffuse until they are transformed to substitutional defects and become immobile in low concentration regions.

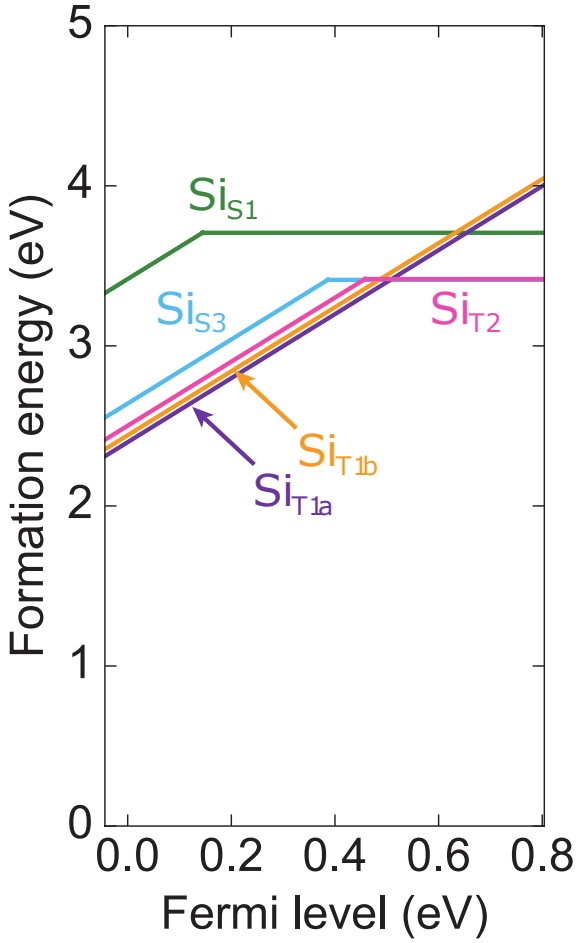


Figure 5.6: Formation energy of Si split interstitials ($S1$ and $S3$) and tetrahedral interstitials ($T1a$, $T1b$, $T2$) in $In_{0.5}Ga_{0.5}As$.

5.4 Conclusions

In this set of studies, we performed Density Functional Theory NEB calculations for all vacancy-assisted and interstitial transitions for Si and intrinsic diffusion in CuAuI-ordered InGaAs. Isolated substitutional Si species are found to be slow diffusers with energy barriers for their vacancy-assisted jumps as high as 2.48 eV. This means that vacancy-assisted Si diffusion in CuAuI-ordered InGaAs is very slow, in agreement with experimental results showing no Si diffusion at low concentration. Our interstitial calculations revealed alternative Si diffusion pathways in CuAuI-ordered InGaAs with low energy barriers. More specifically, split-interstitial configurations are found to be stable and associated transitions to have energy barriers as small as 0.15 eV (compared to the lowest value of 2.12 eV for second-nearest neighbor vacancy-assisted transitions). Such low energy barriers will translate into high diffusivity and could explain the experimentally observed increased diffusivity of Si in InGaAs at high concentration. Si diffusion in real InGaAs systems is bound to be a more complex combination of all the transitions presented in this paper with relative importance depending on growth (which influences the presence or absence of ordered phases), dopant incorporation and annealing conditions.

CHAPTER 6

A CONTINUUM MODEL FOR AMPHOTERIC DOPANT DIFFUSION IN III-VS

6.1 Introduction

It is generally believed that V_{Ga} must play a role in the concentration dependence of Si diffusion in III-Vs. However, the degree to which V_{Ga} vacancies mediate Si diffusion is not fully understood. Also, the nature of the diffusing species is still not fully clear. $Si_{Ga} - Si_{As}$ pairs, Si_{Ga} and Si_{As} have been all individually hypothesized as primary diffusing species in the two leading, competing models for Si diffusion in III-Vs, namely models due to Greiner/Gibbons (GG) [131, 91] and Yu *et al.* [52]. Those two models, which were briefly reviewed in section 3.2.3, make vastly different assumptions around the form in which Si diffuses. Here we discuss limitations of both models and propose an alternative model supported by the new DFT and NEB data garnered in this work.

6.1.1 Limitations of the Greiner and Gibbons model

In the Greiner and Gibbons model, the primary assumptions are as follows:

- $Si_{Ga} - Si_{As}$ is the primary diffusing species
- Si_{Ga} and Si_{As} are immobile when by themselves
- There is a local equilibrium between neutral pairs and their corresponding single donor and acceptor atoms.

Although this model might be correct in assuming a higher concentration of Si-Si pairs at high dopant concentration, there a few limitations within their definition of the model's framework. If single, isolated Si atoms are immobile, it becomes unclear what the mechanism of formation of Si-Si pairs would be. One would imagine that those individual Si atoms would have to move within the crystal lattice until two of them meet and create a Si-Si pair. This model does not incorporate such a possibility, nor does it provide an answer to the question regarding how Si-Si pairs form. We believe that the twin assumptions of immobile single Si atoms and mobile Si pairs are incompatible.

Moreover, our results presented in Chapter 3 show that the single Si atoms are not necessarily immobile, or at least, are not much more immobile than the pairs. Si_{Ga} atomistic jumps to V_{Ga} involve a large energy barrier of 2.19 eV, whereas $Si_{Ga} - Si_{As}$ jumps on the Ga sublattice involve a barrier of 1.84 eV. Therefore, a better assumption would have been that Si complexes are more mobile than individual atoms. Hence, a good diffusion model would have to account for the diffusion of all of those Si species. It is also interesting to note here that our calculations support the assumption that V_{Ga} vacancies are the primary mediators of Si diffusion as jumps involving V_{As} do indeed incur higher energy barriers.

6.1.2 Limitations of the Yu *et al.* model

In Yu *et al.*'s model, the assumptions are as follows:

- $Si_{Ga} - Si_{As}$ pairs are immobile
- Si_{Ga} and Si_{As} are mobile species

- There is a local equilibrium between neutral pairs and their corresponding single donor and acceptor atoms.
- There is a change-over reaction whereby $SiGa$ moves to the As sublattice to become Si_{As} and vice-versa.

The changeover transition suggested in this model is indeed supported by our DFT calculations and involves extremely low energy barriers (0.1 eV from the As to the Ga sublattice, and 0.4 eV for the reverse jump) compared to all the second nearest neighbor jumps (whose energy barriers are high, around 2 eV). However, Yu *et al.*'s assumption that Si pairs are immobile is not supported by DFT calculations. Instead the $Si_{Ga} - Si_{As}$ pairs are found to be more mobile than their single atoms counterparts in stark contradiction with their assumption; thus rendering their model ineffective.

6.2 Continuum model derivation

Since both the GG and Yu *et al.* models have serious limitations and fall short of agreeing with the new computational data acquired in this study, we propose an alternative model supported by our DFT calculations. Without loss of generality, we derive the model using Si as a dopant and GaAs as the III-V material. As mentioned in previous chapters, different group IV atoms (such as Sn and Ge) have similar diffusive behavior as Si, and observations made in GaAs were also made in InGaAs. This model is therefore expected to be applicable to other amphoteric dopants and other III-V compounds.

Our primary assumptions in this new model are as follows:

- $Si_{Ga}-Si_{As}$, Si_{Ga} and Si_{As} can all be mobile species. In other words, *a priori*, we do not exclude any Si species from our model for being immobile.
- $Si_{Ga}-Si_{As}$ pairs are stable species, *i.e.*, once they form, they tend to remain as such. In other words, the equilibrium between Si_{Ga} and Si_{As} species favors $Si_{Ga} - Si_{As}$ pair formation.
- There is a change-over reaction whereby Si_{Ga} can move to the As sublattice to become Si_{As} and vice-versa.

These assumptions are supported by our DFT calculations and address the different limitations of previous models. The underlying idea in this model is that, at low concentration, most Si atoms exist in the form of single Si_{Ga} species, although some Si_{As} species might exist. The vacancy-assisted jumps of individual Si atoms have high energy barriers above 2.0 eV and, therefore, do not lead to significant diffusion. However, at high concentration, other Si species such as $Si_{Ga} - Si_{As}$ pairs and Si split interstitials become present and, since their jumps involve lower energy barriers, lead to the observed enhanced diffusion.

These assumptions are qualitatively consistent with each other and also with experimental observations of compensation effects at high concentration. The fact that Si_{As} and Si_{Ga} individual species are not necessarily immobile added to the existence of the changeover reaction, provides a route for Si pairs to form. Since those pairs are neutral, this mechanism also provides a way for compensation to take place, as Si_{Ga} atoms are *neutralized* by Si_{As} atoms. Moreover, since Si prefers to undergo jumps on the Ga sublattice, it will be often found in the form of $Si_{Ga} - V_{Ga}$ complexes, which provides another mechanism for compensation. We note here that a third mechanism for compensation (or deactivation) is provided by the potential presence of split interstitial species which are un-

likely to contribute to the overall carrier concentration. The relative importance of those three mechanisms is yet to be elucidated.

In terms of primary reactions, our assumptions can be expressed as follows, defining the expression of the change-over reaction (Equation 6.1), the vacancy-assisted transitions (Equation 6.2) and the pair diffusion (Equation 6.3):

$$\mathrm{Si_{As}}^{\alpha_1-} + \mathrm{V_{Ga}}^{\alpha_2-} \rightleftharpoons \alpha_3 \mathrm{e}^- + \mathrm{Si_{Ga}}^{\alpha_4+} + \mathrm{V_{As}}^{\alpha_5+} \tag{6.1}$$

$$\mathrm{Si_{Ga}}^{\alpha_4+} + \mathrm{V_{Ga}}^{\alpha_2-} \rightleftharpoons \mathrm{Si_{Ga}}^{\alpha_4+} + \mathrm{V_{Ga}}^{\alpha_2-} \tag{6.2}$$

$$\mathrm{Si_{As}{-}Si_{Ga}} + \mathrm{V_{Ga}}^{\alpha_2-} \rightleftharpoons \mathrm{Si_{As}{-}Si_{Ga}} + \mathrm{V_{Ga}}^{\alpha_2-} \tag{6.3}$$

In equations 6.1 to 6.3, α_i (with $i = 1, 2, ...$) denote the charge state of the corresponding defects or the number of electrons necessary to charge balance a given reaction. The exact values of these α's depend on the Fermi level as the most thermodynamically stable charge state for each defect varies within the band gap [92]. For the remainder of this chapter, the charge states are omitted for clarity.

Our goal is to derive an expression for the effective diffusivity of an amphoteric dopant (such as Si) in III-Vs. We start by stating the charge balance relationship and the conservation of mass, respectively, given by equations 6.4

and 6.5, where $[Si_{Ga}]$, and $[Si_{As}]$ are the concentrations of silicon residing on Ga and As sites, respectively. $[Si_p]$ and $[Si]_t$ denote the concentrations of $Si_{Ga}-Si_{As}$ pairs and the total dopant concentration, respectively. C_d and C_a are the concentration of donors and acceptors, respectively, in the case of pre-doping, and n and p are the electron and hole concentrations, respectively.

$$[Si]_t = [Si_{Ga}] + [Si_{As}] + 2[Si_p] \tag{6.4}$$

$$p + C_d + [Si_{Ga}] = n + C_a + [Si_{As}] \tag{6.5}$$

Based on the change-over reaction (Eq. 6.1), $[Si_{As}]$ and $[Si_{Ga}]$ are related by the equilibrium constant, k, as follows, where n is the electron concentration:

$$k = \frac{n^{\alpha_3}[V_{As}][Si_{Ga}]}{[V_{Ga}][Si_{As}]} \tag{6.6}$$

We can rewrite Eq. 6.6 as :

$$[Si_{As}] = K[Si_{Ga}] \tag{6.7}$$

where $K = \frac{n^{\alpha_3}[V_{As}]}{k[V_{Ga}]}$

The effective 1-Dimensional diffusion coefficient for Si diffusing in GaAs during an annealing experiment can be described in terms of fluxes by Eq. 6.8. This equation expresses the fact the total flux is the sum of the fluxes of the three Si species Si_{Ga}, Si_{As} and Si_p. The factor of 2 that appears in Eq. 6.4 does not

appear in Eq. 6.8 because we assume the $Si_{Ga} - Si_{As}$ pairs travel together and, as such, can be considered as one species with a given diffusion coefficient, D_p.

$$D^{eff}\frac{\partial[Si]_t}{\partial x} = D_{SG}\frac{\partial[Si_{Ga}]}{\partial x} + D_{SA}\frac{\partial[Si_{As}]}{\partial x} + D_p\frac{\partial[Si_p]}{\partial x} \tag{6.8}$$

where D_{SG} and D_{SA} are, respectively, the diffusivity of Si_{Ga} and Si_{As} species and D_p is the pair diffusion coefficient. Using the mass conservation relation (Eq. 6.4), we can rewrite equation 6.8 as follows:

$$D^{eff} = D_{SG}\frac{\partial[Si_{Ga}]}{\partial[Si]_t} + D_{SA}\frac{\partial[Si_{As}]}{\partial[Si]_t} + D_p\frac{\partial[Si_p]}{\partial[Si]_t} \tag{6.9}$$

$$= D_{SG}\frac{\partial[Si_{Ga}]}{\partial[Si]_t} + D_{SA}\left(1 - \frac{\partial[Si_{Ga}]}{\partial[Si]_t} - 2\frac{\partial[Si_p]}{\partial[Si]_t}\right) + D_p\frac{\partial[Si_p]}{\partial[Si]_t} \tag{6.10}$$

Thus, the effective diffusivity becomes:

$$D^{eff} = \Big(D_{SG} - D_{SA}\Big)\frac{\partial[Si_{Ga}]}{\partial[Si]_t} + \Big(D_p - 2D_{SA}\Big)\frac{\partial[Si_p]}{\partial[Si]_t} + D_{SA} \tag{6.11}$$

Using equations 6.4 and 6.7 we can write:

$$[Si_{Ga}] = \frac{[Si]_t - 2[Si_p]}{1+K} \tag{6.12}$$

Taking the derivative with respect to $[Si]_t$, we obtain:

$$\frac{\partial[Si_{Ga}]}{\partial[Si]_t} = \frac{(1+K)\frac{\partial[([Si]_t - 2[Si_p])]}{\partial[Si]_t} - \big([Si]_t - 2[Si_p]\big)\frac{\partial K}{\partial[Si]_t}}{(1+K)^2} \tag{6.13}$$

which simplifies to:

$$\frac{\partial[Si_{Ga}]}{\partial[Si]_t} = \frac{1}{1+K} - \frac{2}{1+K}\frac{\partial[Si_p]}{\partial[Si]_t} - \frac{[Si]_t - 2[Si_p]}{(1+K)^2}\frac{\partial K}{\partial[Si]_t} \tag{6.14}$$

Inserting the expression for $\frac{\partial[Si_{Ga}]}{\partial[Si]_t}$ given by Eq. 6.14 in Eq. 6.11 leads to:

$$\begin{aligned} D^{eff} =& \left(D_{SG} - D_{SA}\right)\left[\frac{1}{1+K} - \frac{2}{1+K}\frac{\partial[Si_p]}{\partial[Si]_t} - \frac{[Si]_t - 2[Si_p]}{(1+K)^2}\frac{\partial K}{\partial[Si]_t}\right] \\ &+ \left(D_p - 2D_{SA}\right)\frac{\partial[Si_p]}{\partial[Si]_t} + D_{SA} \end{aligned} \tag{6.15}$$

After simplification, we finally obtain:

$$\begin{aligned} D^{eff} =& \frac{D_{SG} + KD_{SA}}{1+K} - \frac{D_{SG} - D_{SA}}{(1+K)^2}\left([Si]_t - 2[Si_p]\right)\frac{\partial K}{\partial[Si]_t} \\ &+ \left[D_p - \frac{2(D_{SG} + KD_{SA})}{1+K}\right]\frac{\partial[Si_p]}{\partial[Si]_t} \end{aligned} \tag{6.16}$$

We now seek to derive an expression for $[Si_p]$ as a function of $[Si]_t$. We can define an equilibrium constant between single Si atoms and Si pairs as follows, accounting from the fact that two Si atoms would have to react to create an Si-Si pair:

$$[Si_{Ga}]^2 = K_p[Si_p] \tag{6.17}$$

which can be rewritten as:

$$[Si_{Ga}] = Kp^{1/2}[Si_p]^{1/2} \tag{6.18}$$

The exact reaction responsible for pair formation remains an open question, but is not consequential for the remainder of this derivation.

Now using mass conservation (Eq. 6.4) and Equation 6.7 we can write:

$$[Si]_t = [Si_{Ga}] + [Si_{As}] + 2[Si_p] \tag{6.19}$$

$$= (1 + K)[Si_{Ga}] + 2[Si_p] \tag{6.20}$$

$$= (1 + K)K_p^{1/2}[Si_p]^{1/2} + 2[Si_p] \tag{6.21}$$

We can then solve for $[Si_p]$ as a function of $[Si]_t$ to obtain:

$$[Si_p] = \left(\frac{-\alpha + \sqrt{\alpha^2 + 8[Si]_t}}{4}\right)^2 \tag{6.22}$$

where $\alpha = (1 + K)\sqrt{K_p}$.

K and K_p are the primary tuning parameters in this model and they encode contributions of the Fermi level and the relative concentrations of V_{Ga} and V_{As} as determined by experimental conditions (*e.g.,* arsenic overpressure is known to influence the $[V_{Ga}]/[V_{As}]$ ratio).

The derivative of $[Si_p]$ with respect to $[Si]_t$ is:

$$\frac{\partial[Si_p]}{\partial[Si]_t} = \frac{1}{2}\left(1 - \frac{\alpha}{\sqrt{\alpha^2 + 8[Si]_t}}\right) \tag{6.23}$$

We can then write the final expression for the effective diffusivity as a function of concentration as:

$$
\begin{aligned}
D^{eff} =& \frac{D_{SG} + KD_{SA}}{1+K} - \frac{D_{SG} - D_{SA}}{(1+K)^2}\Big([Si]_t - 2[Si_p]\Big)\frac{\partial K}{\partial [Si]_t} \\
&+ \frac{1}{2}\left[D_p + \frac{2(D_{SG} + KD_{SA})}{1+K}\right]\left(1 - \frac{\alpha}{\sqrt{\alpha^2 + 8[Si]_t}}\right)
\end{aligned}
\tag{6.24}
$$

In more general terms, the effective diffusivity of an amphoteric dopant A in a III-V semiconductor can be expressed by Eq. 6.25 where D_{III} and D_V are, respectively, the diffusivity on the III and V sublattices, K the compensation ratio, $[A]_t$ the total dopant concentration, $[A_p]$ the concentration of dopant pairs that could form at high concentration and α a Fermi level-dependent parameter that controls the fraction of $[A]_t$ that exist in the form of $[A_p]$.

$$
\begin{aligned}
D^{eff} =& \frac{D_{III} + KD_V}{1+K} - \frac{D_{III} - D_V}{(1+K)^2}\Big([A]_t - 2[A_p]\Big)\frac{\partial K}{\partial [A]_t} \\
&+ \frac{1}{2}\left[D_c + \frac{2(D_{III} + KD_V)}{1+K}\right]\left(1 - \frac{\alpha}{\sqrt{\alpha^2 + 8[A]_t}}\right)
\end{aligned}
\tag{6.25}
$$

6.3 Results and Discussion

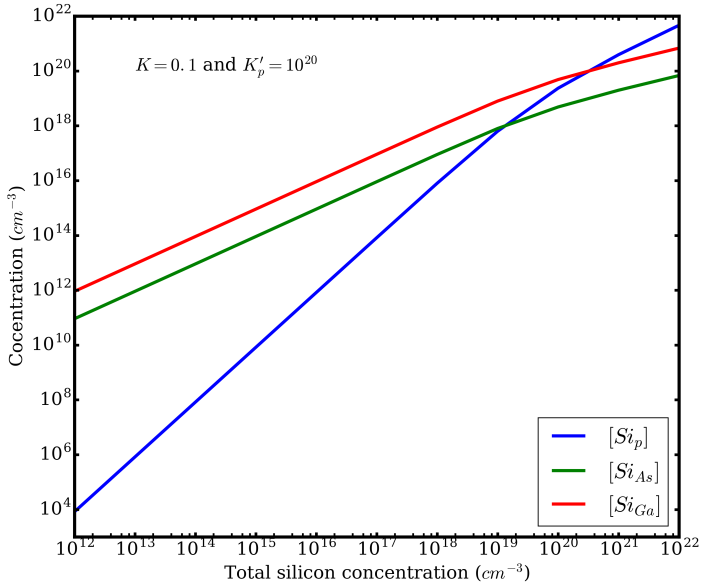
6.3.1 Concentration of Si species: An example in GaAs

To plot the concentration of the different Si species, we assume a constant value for K. In general, for n-doped GaAs, it is expected that the concentration of Si_{As} will be much less than the concentration of Si_{Ga}. We therefore arbitrarily use a value of 0.1 for K. In Fig. 6.1, the concentration of Si pairs, Si_{Ga} and Si_{As} are plotted as a function of total Si concentration for different values of K_p. At low Si concentration, most of the Si atoms will occupy Ga substitutional sites with a small fraction occupying As sites. The concentration of Si pairs will be

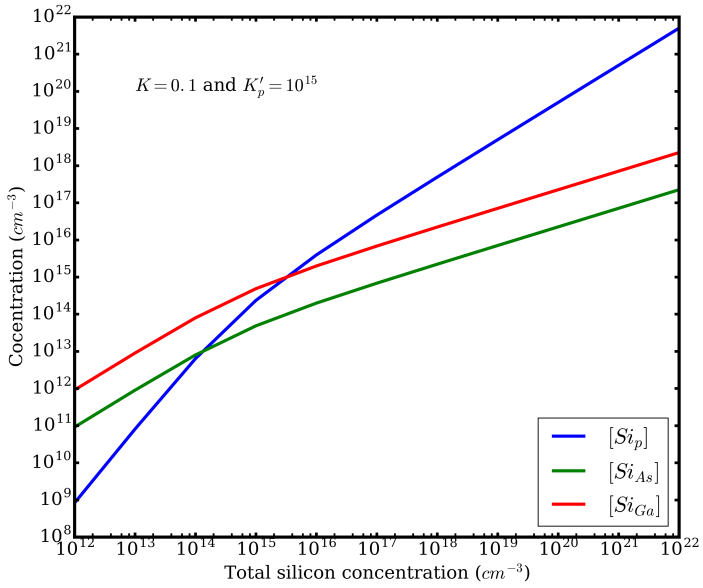
negligible. At high Si concentration, the majority of the population of Si atoms exist in the form $Si - Si$ pairs (or potentially higher order complexes). The two concentrations equal each other at a value determined by the parameter K_p'. This is in complete agreement with the assumptions of our model.

6.3.2 Effective Si diffusivity in GaAs

We now apply our model to the diffusion of Si in GaAs and compare the results to those in the literature. To calculate an effective silicon diffusivity, we require values for D_{SG}, D_{SA} and D_p. We assume an Arrhenius expression for the diffusion coefficients as a function of temperature as given by Eq. 6.26, where D_0 is the exponential prefactor, Q, the activation energy which includes contributions from the enthalpy of formation of both the Ga vacancies and the Si substitutional defects and their migration energies, E_m. It is worth noting here that the DFT calculations performed in this work give access only to the migration energy. We further assume that only a small fraction of Si atoms exist in the form of isolated Si_{As} species (*i.e.*, $K \ll 1$), and that D_{SA} has a negligible contribution to the effective diffusivity, as suggested by our DFT calculations. These additional assumptions lead to the simplified expression given by Eq. 6.27. We find optimal values for K_p' as well as the prefactors and activation energy for both D_{SG} and D_P that minimize the error with experimental data points for GaAs. Results of this fitting process are given in table 6.1 and Fig. 6.2 for the diffusivity at 1323 K and 1273 K in two distinct experiments (diffusion data extracted from publication by Yu *et al.* [52]).



(a)



(b)

Figure 6.1: Concentration of the different species of Si atoms of interest in our continuum model of diffusion as a function of total Si concentration and the parameters K and K_p. The influence of K_p on the relative distribution of Si atoms into Si_{Ga}, Si_{As} and Si pairs is significant.

$$D = D_0 \exp(-Q/kT) \tag{6.26}$$

$$D^{eff} = D_{SG} + \frac{1}{2}(D_p - 2D_{SG})\left(1 - \frac{\alpha}{\sqrt{\alpha^2 + 8[Si]_t}}\right) \tag{6.27}$$

Although the fitting obtained with constant α values is reasonable (see Fig. 6.2), it can be greatly improved by incorporating its concentration-dependence in our calculations. Multiple published *ab initio* data for $GaAs$ and $InGaAs$ [92, 114] have, in fact, shown a decrease of V_{Ga} and V_{In} formation energies with Fermi level and dopant concentration. Therefore, we can reasonably assume that K_p and consequently α would change with dopant concentration. The exact form of this dependence is unknown, but we can assume a sigmoid-like dependence (Eq. 6.28) whereby pair formation is less likely at low concentration and then transitions to a high probability at high concentration. Such an assumption is in line with formation energy calculations in $InGaAs$ [92] that show an increase in $[V_{III}]$ concentration at high dopant concentration. A, B, Q and $[Si]^*$ in Eq. 6.28 are fitting parameters defining the minimum, maximum and scale of changes of K_p with Si concentration.

$$K_p = A + \frac{B - A}{1 + Q \exp(-[Si]_t/[Si]^*)} \tag{6.28}$$

Using Eq. 6.28, we now obtain improved agreement between our model and the experimental diffusivity data for all temperatures as shown in Fig. 6.3 (diffusion data extracted from publication by Yu *al* [52]). Values for all the fitting parameters are given in Table 6.2.

Parameter	Value for GaAs
K_p	5.89×10^{20}
D_0^{SG}	4.27×10^{-4}
Q^{SG}	2.21
D_0^P	3.82×10^2
Q^P	3.22

Table 6.1: List of parameter values obtained from fitting our diffusion model to experimental data for Si diffusion in GaAs assuming constant α values.

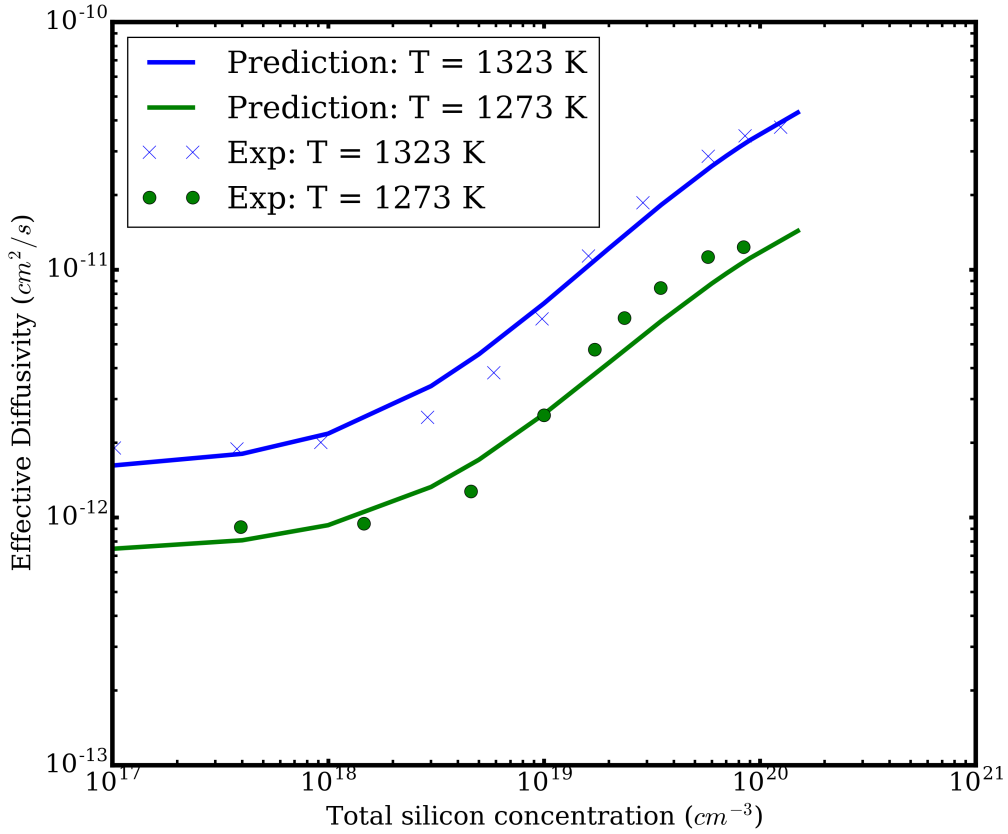


Figure 6.2: Our effective diffusivity model fitted to experimental data for Si diffusion in GaAs for constant α values

Parameter	1323 K	1273 K	1253 K	1213 K
A	1.72×10^{21}	2.74×10^{21}	7.37×10^{24}	3.0×10^{25}
B	3.895×10^{20}	5.63×10^{20}	1.4×10^{24}	4.63×10^{24}
Q	10	10	10	10
$[Si]^* \, [cm^{-3}]$	2.65×10^{18}	2.23×10^{18}	1.29×10^{18}	1.65×10^{18}
$D_{SG} \, [cm^2/s]$	1.76×10^{-12}	9.13×10^{-13}	5.55×10^{-13}	3.26×10^{-13}
$D_p \, [cm^2/s]$	3.29×10^{-10}	1.54×10^{-10}	1.47×10^{-7}	3.02×10^{-7}

Table 6.2: List of parameter values obtained from fitting our diffusion model to experimental data for Si diffusion in GaAs assuming concentration-dependent α values. Again we assume $K \to 0$.

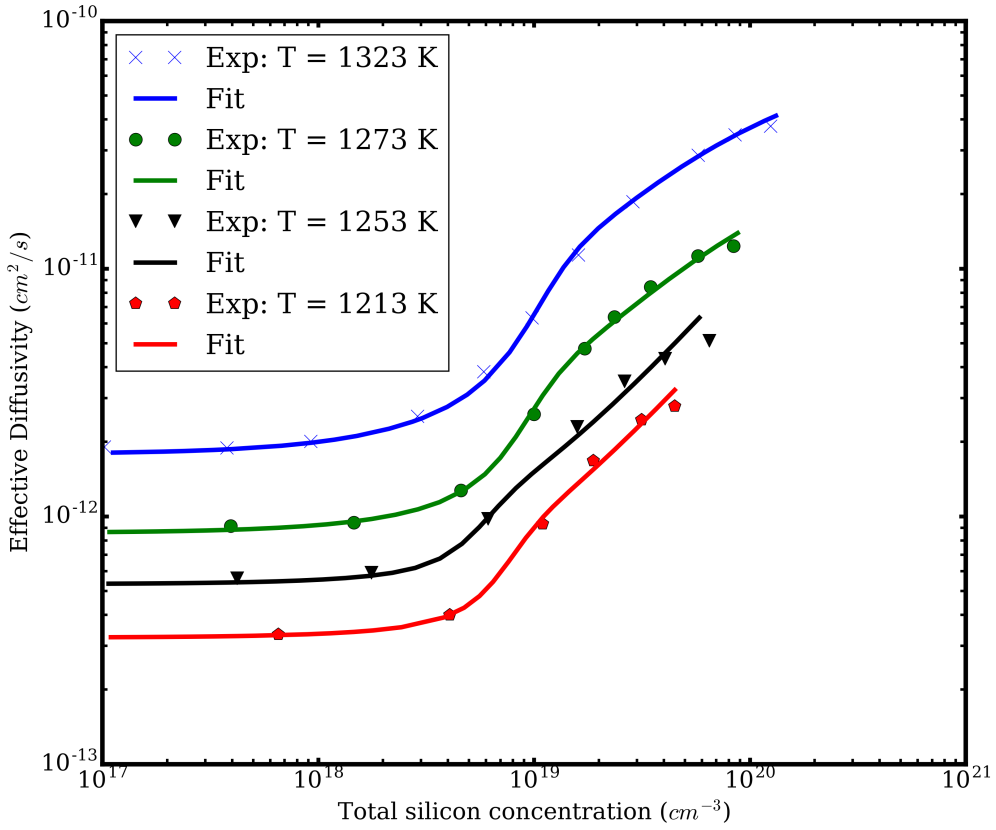


Figure 6.3: Our effective diffusivity model fitted to experimental data for Si diffusion in GaAs with concentration-dependent α values.

6.3.3 Comparison with previous models

The effective diffusivity obtained by Greiner and Gibbons in their model is given by Eq. 6.29, where D_p is the pair diffusion coefficient, C_t the total silicon concentration, and K_p the pair equilibrium constant. This expression is plotted in Fig. 6.4 for an arbitrary value of $D_p = 10^{-12}$ cm^2/s. As shown in Fig. 6.4, the expression leads to a linear dependence of the log of the effective diffusivity, $ln(D_{eff})$, on total Si concentration at low concentration which saturates to D_p at high concentration. Such a linear relationship is not supported by experimental data. In contrast, the expression developed for D_{eff} in this work is able to predict a constant low diffusivity at low silicon concentration and a high diffusivity at high concentration, in agreement with experiments.

$$D_{eff} = D_p[1 - (1 + 2C_t/K_p)^{-1/2}] \tag{6.29}$$

On the other hand, although our assumptions and final analytical expression are different from Yu *et al.*'s model, surprisingly, the qualitative dependence of D_{eff} on Si concentration obtained is similar. Both models predict a constant, low effective diffusivity at low silicon concentration, and a high diffusivity at high Si concentration. The reduced expression used by Yu *et al.* to calculate D_{eff} is given by Eq. 6.30 where $\gamma = [Si_{As}]/[Si_{Ga}]$ is a concentration-dependent compensation ratio. This expression leads to $D_{eff} = D_{SG}$ for $\gamma = 0$, and $D_{eff} = (D_{SG} + D_{SA})/2$, which is a simple average of the diffusivity of Si_{Ga} and Si_{As} species for $\gamma = 1$. As discussed in this chapter, the assumption of high diffusivity for Si_{Ga} and Si_{As}, with no considerations of complexes, is not supported by our DFT calculations. Moreover, to fit their model to experiments, Yu *et al.* had to employ very high values for γ at high silicon concentration. Such

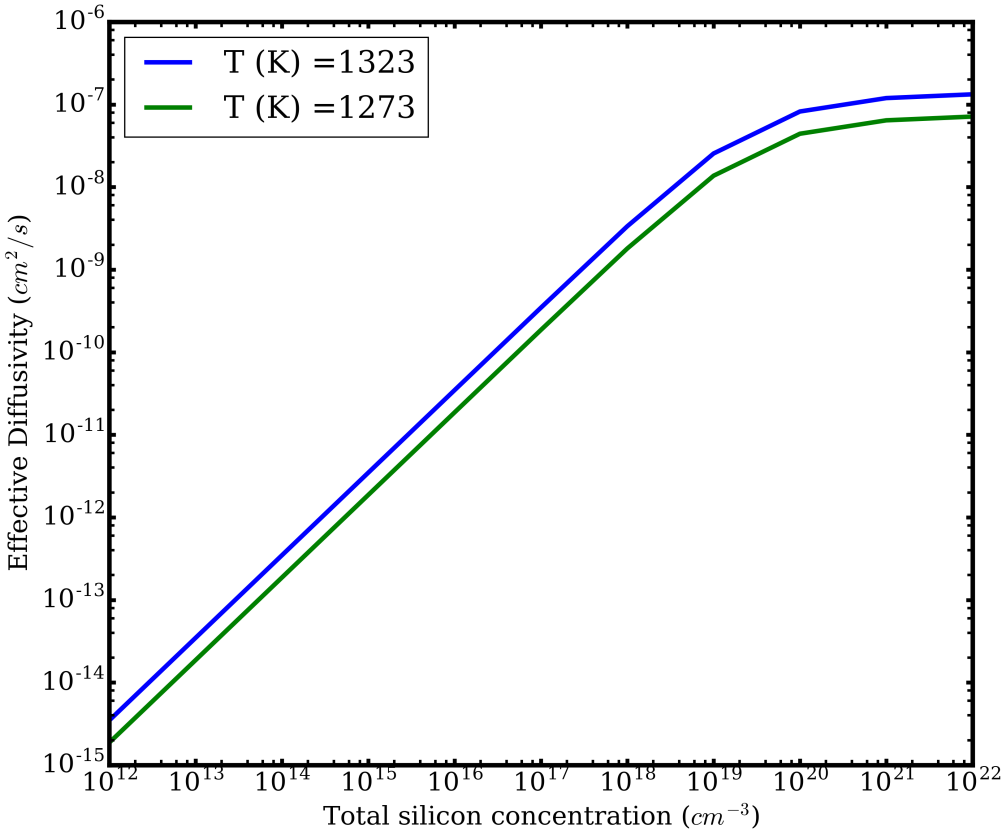


Figure 6.4: Effective diffusivity predicted by the Greiner model as a function of total Si concentration. This model is unable to capture the sigmoid-like shape of experimentally observed effective diffusivity as a function of dopant concentration.

high values for γ ($\gamma \rightarrow 1$) are questionable as this would mean that almost all the Si atoms occupy As sites in an n-type doping regime, which is unphysical.

$$D_{eff} = (D_{SG} + \gamma D_{SA})/(1 + \gamma) \tag{6.30}$$

6.3.4 Diffusion profile in InGaAs

We show in Fig. 6.5 that our model can successfully predict the diffusion profiles obtained for Si in InGaAs. The previous model used to fit such highly unusual diffusion profiles was based on an effective diffusivity given by $D_{eff} = D_f + D_c\left(\frac{C}{10^{19}}\right)^n$, where D_f, D_c and n were fitting parameters [1]. Such an expression for the effective diffusivity does not seem to have any physical basis and may have been proposed arbitrarily. Moreover, it leads to diffusivity values that grow to infinity for high dopant concentration. Such a behavior is hardly physical. In contrast, our model behaves well at high dopant concentration and has sound thermodynamic and kinetic basis for its derivation. Values used for the fitting process are given in Table 6.3.

6.3.5 Limitations of the Model

Surface states

One key difference between our model and that of Yu *et al.* lies in the inclusion of surface states in the latter. In Yu *et al.*'s treatment, surface states, mainly corresponding to deep acceptor states, are taken into account to explain an apparent

Parameter	Value for InGaAs
A	4.18×10^{23}
B	2.69×10^{21}
Q	10
$[Si]^*$	5.89×10^{20}
D_0^{SG}	7.89×10^4
Q^{SG}	2.96
D_0^P	5.9×10^6
Q^P	2.61

Table 6.3: List of parameter values used to fit our diffusion model to experimental data for Si diffusion in InGaAs at 750 C.

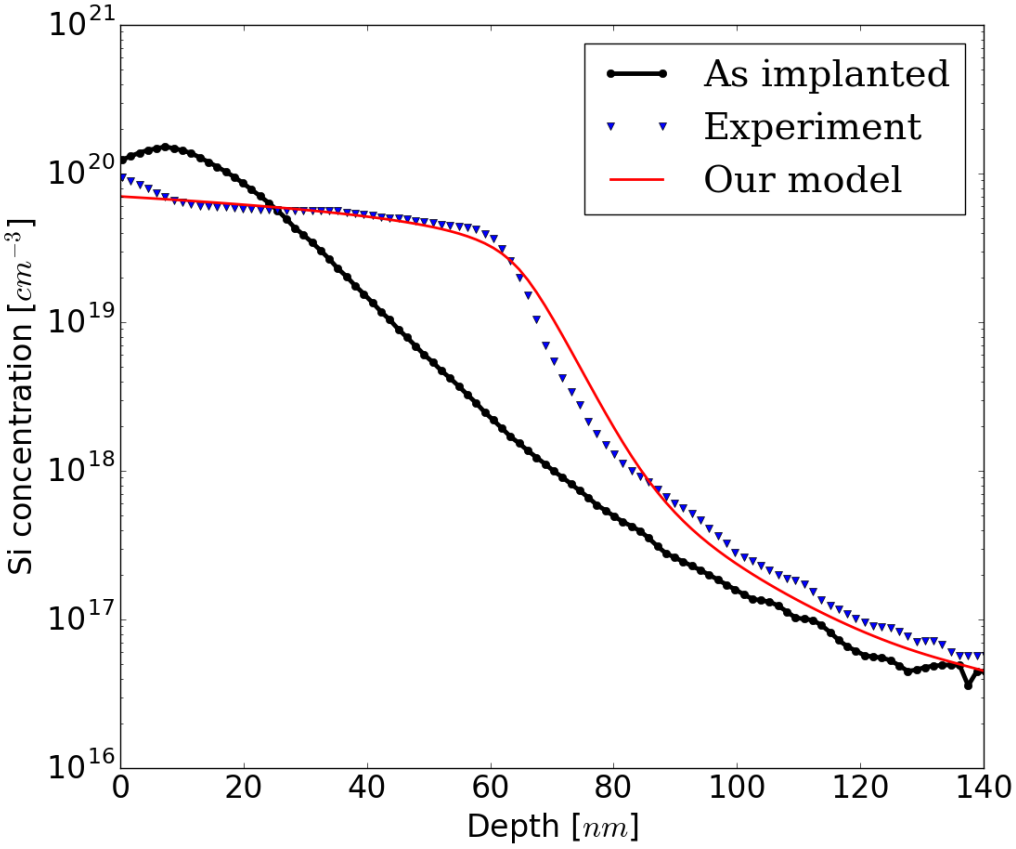


Figure 6.5: Fitting of experimental Si diffusion profile in InGaAs at 750 C with our model. Data from Aldridge Jr. *et al.* [1].

decrease in effective diffusivity attributed to a decrease in free electrons. Such considerations are not included in our model as it is more targeted towards diffusion within the bulk crystal. However, a good understanding of the effects of surface states could be very useful for diffusion experiments that maintain a constant concentration of dopants at the surface. Inclusion of such effects in future iterations of this model could therefore be favorably considered.

Split interstitials

Our DFT calculations have shown that Si split-interstitial species are very fast diffusers in all of the three III-V semiconductors we studied (GaAs, InAs and InGaAs). Therefore, in conditions that promote their formation, one can expect to see a significant increase in diffusion. The exact nature of such conditions are not yet clear and experimental evidence of their presence has not been gathered at present. We think ion implantation could be a potential method that leads to the creation of such Si split interstitial species. In either case, their presence would lead to enhanced diffusion as they move quickly through the crystal lattice until they reach vacancies to annihilate. Since we don't expect those species to be present in conditions such as diffusion from a constant surface concentration, they were not included in our model. However, their addition to this model would be desirable in the future for greater applicability.

6.4 Conclusions

In conclusion, the DFT and NEB calculations performed in this work are used to inform the development of a new general continuum model applicable to

amphoteric diffusion in III-Vs. Our model assumes that diffusion of complexes, such as Si pairs, dominate at high dopant concentration and that diffusion of isolated Si atoms dominate at low concentration. We compare our new model to the two most prominent previous models and show how this new model is an improvement. We also validate our model by showing that it can successfully capture the experimentally observed concentration-dependence of Si diffusion in GaAs.

CHAPTER 7

MOLECULAR FINGERPRINTS FOR MACHINE LEARNING APPLICATIONS

7.1 Introduction

"Feature extraction," or "feature engineering," is a key step in machine learning studies which helps ensure that accurately predictive models can be built from available data. In this context, feature engineering refers to the process of identifying, evaluating and applying one or more transformations (usually mathematical operations) to a dataset to enhance its suitability for machine learning use. For applications involving prediction of properties based on molecular structures, this process usually leads to the computation of a vector or matrix called a "fingerprint," or "descriptor," that represents a chemical structure or environment (See Fig. 7.2). Beyond machine learning applications, fingerprints obtained in this manner can also be used for a systematic study of structure-property relationships, including structural analysis and structural comparison for the identification of configurations of interest and for property prediction. [78]

To be suitable for machine learning studies, especially forces and energy predictions, fingerprints have to satisfy a number of requirements such as invariance under rotation, translation and permutation of atoms of the same chemical nature. [133, 78, 134] Differentiability and speed of computation are also

This chapter was published in Molecular Systems Design and Engineering [132]

desirable. [78, 134] The need for symmetrization of atomic coordinates to satisfy the requirement of invariance under permutation of chemically identical atoms, especially for use in Neural Networks, has been recognized for at least two decades.[135] This need arises because Cartesian coordinates, traditionally used to represent molecular systems, are not invariant under those geometric operations. For example, permutation of two atoms of the same type does not affect the properties of molecular systems, but it changes their Cartesian coordinates. Hence the need arises for alternative representations that are permutation-invariant. Using the same argument, one can conclude that rotation and translation-invariance are also necessary. Differentiability is usually desirable for force predictions.

7.2 Overview ot fingerprinting strategies

Many different fingerprints have been proposed over the years for machine learning applications. Some examples include Coulomb Matrix,[136] Bag of Bonds,[137] symmetry functions, [138] bispectrum, [139] Zernike, [134] and others. More recently, Artrith *et al.* proposed a way to generate fingerprints for compositions containing many species.[140] A method to automatically find appropriate descriptors was suggested by Ghiringhelli *et al.*[141] A fingerprint called a "depth map" (D-map) that readily inverts into molecular geometries was tested by Yao *et al.* [142] Other descriptors such as Bonds-in-Molecule [143] and many-body expansion approaches have also been recently explored. Sadeghi *et al.* evaluated a number of fingerprints of the form $n \times n$ matrices, where n is the number of atoms in the system,[144] including contact [145] and overlap [146] matrices, useful for evaluating similarity between structures.

Since there are so many different fingerprints proposed in the literature, we propose a fingerprinting categorization scheme that affords an easier and more systematic analysis.

A number of other fingerprints exist in the field of drug design to which the larger materials science community probably owes the term "molecular fingerprint." [147, 148, 149] Those molecular representations, for the most part, were created years before the relatively new fingerprints addressed in this paper and were specifically designed for comparisons and quantitative structure-activity relationship studies of molecules relevant to biological applications. Notwithstanding their usefulness otherwise, most of those fingerprints are not particularly suitable for spatially resolved machine learning studies of molecular systems as they usually do not satisfy the requirements discussed above. To streamline the process of using machine learning in computational materials science, we have developed a codebase called SEING (an old French word for signature) which is described in more detail in Section 7.5.

7.3 Classification of fingerprints

In Table 7.1, we list major fingerprints in the current literature designated by type and the year in which they were first proposed, covering the past ten years. From a high-level perspective, fingerprints can be divided into atom-centered (or local) and global descriptors, based on whether they represent the environment around a given atom or are constructed to represent an entire molecule or crystal (see Fig. 7.1). Such a differentiation has already been used previously in the literature. [134, 150, 78] Atom-centered fingerprints are particularly suit-

able to build machine learning-based force fields and for the prediction of local properties. Global fingerprints are, by design, more suitable for the prediction of macroscopic properties.

Atom-centered fingerprints can be further divided into whether they are based on symmetry functions, basis set decomposition or similarity measures. Symmetry function-based representations use summations over bonds, angles or dihedrals of the same nature, or summation over functions thereof, for all atoms within a given sphere around the atom of interest. Fingerprints that fall within this category include Behler-Parrinello, [138, 76] and Jose-Artrith-Behler. [151] They usually differ by the functional forms used in the summation. Basis set decompositions, on the other hand, rely on a local atomic density function defined around the atom of interest which is then expressed as a linear combination of a suitable basis set. Subsequent operations are usually used to ensure invariance under rotation and/or translation, and so on. Examples of such fingerprints include Bispectrum, [139] Zernike, [134] and the Artrith-Urban-Ceder method, which was specially developed to support multi-component systems. [140] The Smooth Overlap of Atomic Positions (SOA) [152] and Graph-Approximated Energy (GRAPE)[153] methods fall into a third category in which local environments are compared directly with the help of an appropriate kernel function without the explicit computation of a fingerprint.

Global fingerprints, on the other hand, can be divided into ones that involve only pair distances and ones that include higher-order interactions. The former subcategory has the largest number of fingerprints and includes extended connectivity, [154] Coulomb Matrix, [136] as well as Contact, [145] Overlap, [146] Hessian and Hamiltonian matrices, [144] Partial Radial Distribution Functions

Table 7.1: Overview of the main fingerprinting strategies for machine learning applications proposed over the past ten years using the classification scheme developed in this paper.

Fingerprints	Classification	2007-2010	2011-2013	2014-2016	2017
Atom-Centered	Symmetry Functions	Behler-Parinello [138]	Behler[76] Jose-Artrith-Behler [151]	AGNI [158]	AGNI2.0 [159] wACSF [160]
	Basis Set Decomposition	Bispectrum[139]		Zernike[134]	Artrith-Urban-Ceder [140]
	Kernel-Based		SOAP [152]		GRAPE[153]
Global	Pair Distances only	Extended Connectivity[154]	Coulomb Matrix[136] Contact Matrix[145] Overlap Matrix[146] Hamiltonian Matrix[144] Hessian Matrix[144]	PRDF[155] Bag of Bonds[137] Fourier Series[156]	Bonds in Molecule[143]
	Distances, Angles, Dihedrals			BAML[161]	Many-Body Expansion[142] MBTR[162] MARAD[70] HDAD[70]
	Graph-Based			Molecular Graph[163]	Property-Labeled[164] Materials Fragments
	Others		ALF[165]	Motif-Based[166]	D-Map[142] RAC [167]

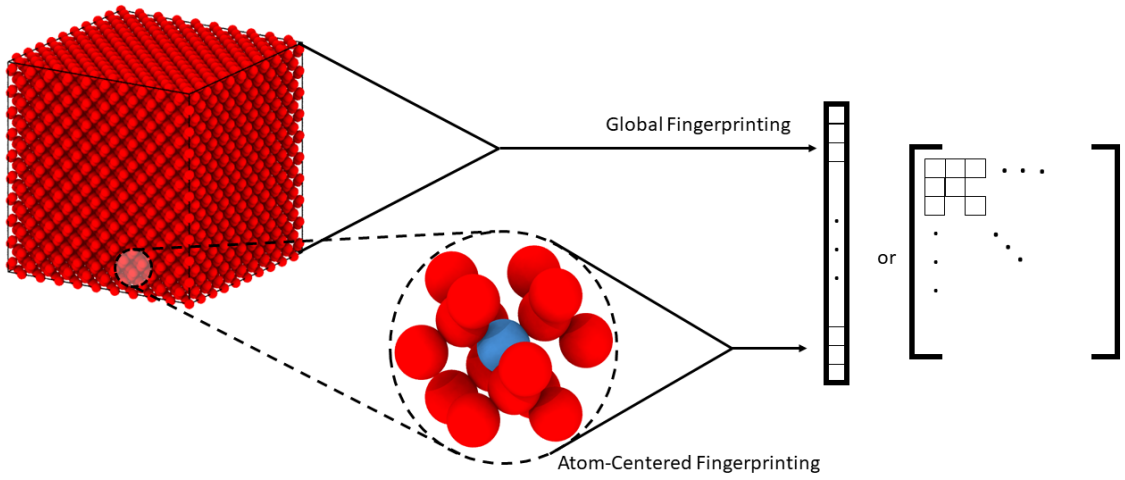


Figure 7.1: Two distinct strategies exist to generate fingerpints: global and atom-centered. The former is generally suitable for macroscopic property predictions whereas the latter can be used for local property predictions and are particularly relevant to force field development.

(PRDF), [155] Bag-of-Bonds, [137] Fourier Series decomposition of Radial Distribution Functions, [156] Bonds-in-Molecule, [143] connectivity count and encoded distances. [157] Those fingerprints are built using different schemes or functions of some or all pair distances in the molecule or crystal of interest. Sometimes, the fingerprints also use information related to the nature of the different elements, such as partial charges, atomic numbers, *etc.*.

The latter subcategory includes recently developed fingerprints such as Bond Angle Machine Learning (BAML), [161] Many-Body Expansion (MBE), [142] Many-Body Tensor Representation, [162] Molecular atomic radial angular distribution (MARAD), [70] and the Histogram of Distances, Angles and Dihedral angles (HDAD). [70] Those fingerprints include higher-order interaction terms such as angles and dihedrals and usually lead to more accurate predictions compared to their pair distances-only counterparts.

A third subcategory of global fingerprints is comprised of graph-based descriptors, such as Molecular Graphs [163] and Property-Labelled Materials Fragments, [164] which are suitable for use in graph-based convolutional Neural Networks. Some other descriptors can be included in their own subcategory; these include the Atomic Local Frame (ALF)[165], motif-based fingerprint [166] and the Revised Autocorrelation (RAC) fingerprints [167] which are inspired by descriptors used in cheminformatics. This subcategory would also include the D-map [142] fingerprint that can be readily mapped back to chemical structures.

In this section, we have proposed a cataloging scheme for fingerprints for machine learning applications. It is important to note that the fingerprints mentioned in this paper are primarily based on the spatial coordinates of atoms in the molecular system of interest. This essentially assumes that such coordinates

must have been acquired through means such as *ab initio* calculations, prior to performing the machine learning studies. Moreover, though they are usually designed to include all atoms in a molecular system, global fingerprints can sometimes be used as local fingerprints by including only atoms that fall within a given cutoff distance of a central atom of interest.

A whole other class of fingerprints can be constructed based solely on topological or connectivity information and/or nature of the chemical species and chemical motifs (such as fragments, rings, *etc.*,) in the structure. Such fingerprints do not explicitly encode spatial coordinates of atoms but can successfully be used to predict various properties using machine learning techniques. [168, 71, 81] A number of the fingerprints from cheminformatics (which are excluded from this study) would also fall into that class. While the categorization scheme suggested here is arbitrary, it provides a suitable framework to better understand and analyze different fingerprints proposed in the literature.

7.4 Review of key fingerprints

7.4.1 Symmetry Function Fingerprints

In general, symmetrization strategies involve a direct summation over pairwise distances and/or angles between three atoms or a summation over parametric functions of those distances and angles. These "symmetry functions" have the effect of making atoms lose their individuality to all other symmetrically and chemically equivalent atoms. Examples include the atom-centered G^2 and G^4 functions proposed by Behler and Parrinello (BP) [138, 76] and given by Equa-

tions 7.1 and 7.2, respectively. For a given atom i, the BP method uses distances R_{ij} to all atoms j within a cutoff distance R_c, as well as angles θ_{ijk} centered on atom i and involving atoms j and k with $j \neq k$. The cutoff function, f_c, ensures a smooth transition to a contribution of zero for atoms that are outside of the R_c distance. An example of a cutoff function is given in Equation 7.3. Generating a finite number of those symmetry functions based on a set of predefined parameter values corresponding to the width η and center R_s of the Gaussians in G^2 as well as λ, η and ζ of the Gaussians in G^4 ensures fingerprints are an effective representation of the chemical environments that they seek to capture. The number of parameters to use, as well as their specific values defining the spatial resolution achieved, becomes a design decision left to the discretion of the modeler.

$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}) \tag{7.1}$$

$$G_i^4 = 2^{1-\zeta} \sum_{j,k\neq i} (1+\lambda\cos\theta_{ijk})^\zeta e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)^2} f_c(R_{ij}) f_c(R_{ij}) f_c(R_{ij}) \tag{7.2}$$

$$f_c = \begin{cases} 0.5[\cos(\frac{\pi R_{ij}}{R_c}) + 1] & \text{for} \quad R_{ij} \leq R_c \\ 0 & \text{for} \quad R_{ij} > R_c \end{cases} \tag{7.3}$$

Other types of symmetry and cutoff functions have been proposed by different authors. For example, pair-centered symmetry functions have been used by Jose *et al.* [151] to construct Neural Network potentials. Another approach, based on Permutation-Invariant Polynomials, has been suggested by Jiang *et*

al.[169] and by Li *et al.* [170] In this method, Cartesian coordinates are replaced by a summation over symmetrized Morse-like monomials that include all possible nuclear permutations in the system. AGNI [158, 159] fingerprints are another example of symmetry function-based fingerprints developed to predict atomic forces directly. As shown by Equation 7.4, AGNI fingerprints are very similar to the G^2 components of the BP fingerprints (see Eq. 7.1). However, they include a direction-specific coefficient given by the ratio of the α component (where α=x, y or z) of the pair distance, r_{ij}^{α}, divided by the pair distance r_{ij} between atoms i and j, akin to a derivative of the BP fingerprints in that direction. In addition, especially in their newer version of AGNI, those authors showed that use of Gaussian centers, a_k (labeled R_s in G^2), as the main parameter, rather than the width, w (labeled η in G^2) of the Gaussians, led to better predictive capabilities for the machine learning models.

$$V_{i,\alpha,k} = \sum_{j \neq i} \frac{r_{ij}^{\alpha}}{r_{ij}} \frac{1}{\sqrt{2\pi w}} e^{-0.5(\frac{r_{ij}-a_k}{w})^2} f_c(r_{ij}) \tag{7.4}$$

Although symmetry functions have been successfully used for a number of applications, one significant drawback of this approach has been the size increase of the descriptors with the number of elements. Traditionally, a new set of radial functions are created for each atom type and a new set of angular functions for every combination of elements, leading to a significant increase of fingerprint size with number of species. This usually makes it impractical to use symmetry functions for systems containing more than four elements.[78] However, the recent Weighted Atom-Centered Symmetry Functions (wACSF)[160] method proposed by Gastegger *et al.* attributes a species-specific weight to each term in the sum for G^{rad} and G^{ang}, hence avoiding the fingerprint dimensional-

ity increase with number of elements. A similar approach was used by Artrith *et al.*. [140]

7.4.2 Bispectrum Fingerprints

The bispectrum fingerprinting approach, previously used for image or pattern recognition,[171] was first proposed by Bartók *et al.*[139] for the representation of molecular structures. It relies on the atomic density around a center atom, i, as defined by Equation 7.5 where $f_{cut} = 1/2 + \cos(\pi r/r_{cut})$ and δ is the Dirac delta function. The local atomic density (LAD) is invariant to permutation by construction. It is further projected onto the four-dimensional unit sphere and expressed in terms of 4D spherical harmonics. The bispectrum components, $B_{j_1,j_2,j}$, are then built from the coefficients, $c^j_{m'm}$, of the expansion according to Equation 7.6, where $C^{jm}_{j_1m_1j_2m_2}$ are Clebsch-Gordan coefficients and $j, j_1, j_2 \leq J_{max}$ with J_{max} being the only parameter. This additional transformation ensures invariance of the fingerprint to rotation and translation.

$$\rho_i(\boldsymbol{r}) = \delta(\boldsymbol{r}) + \sum_j \delta(\boldsymbol{r} - \boldsymbol{r}_{ij}) f_{cut}(|\boldsymbol{r}|) \tag{7.5}$$

$$B_{j_1,j_2,j} = \sum_{m'_1,m_1=-j_1}^{j_1} \sum_{m'_2,m_2=-j_2}^{j_1} \sum_{m',m=-j}^{j} c^j_{m'm} C^{jm}_{j_1m_1j_2m_2} C^{jm'}_{j_1m'_1j_2m'_2} c^{j_1}_{m'_1m_1} c^{j_2}_{m'_2m_2} \tag{7.6}$$

This method has two main advantages over symmetry functions. First, it eliminates the need to choose a number of parameters defining the spatial resolution of the fingerprints while increasing the fingerprint size, as in the case

of symmetry functions and, second, it can be systematically improved with the addition of more spherical harmonics in the summation (as given by J_{max}). This method can also be easily extended to multi-component systems using a factor w_j for the contribution of each atom in the LAD. In their original paper, Bartók *et al.* used Gaussian Process Regression with the bispectrum representations to predict a number of properties for carbon, silicon and germanium crystals. [139] They called their resulting model a Gaussian Approximation Potential (GAP), which differs from another bispectrum-based approach by Thompson *et al.* called the Spectral Neighbor Analysis Method (SNAP) in which atom energies are assumed to be linearly dependent on the bispectrum components. [172] This new assumption afforded easier fitting of the resulting potential. However, a direct and comprehensive comparison between those two methods has yet not been performed.

7.4.3 Zernike Fingerprints

Zernike moments, originally used for 3D shape retrieval in the machine learning community,[173] were recently proposed by Khorshidi *et al.* to build molecular fingerprints.[134] In their approach, a LAD function similar to Equation 7.5 is defined and expanded with 3D Zernike basis functions. Zernike basis functions are defined as products of Zernike polynomials (which are radial basis functions defined inside the unit sphere) and spherical harmonics (which form a basis set on the surface of the unit sphere). This approach is similar to the construction of the bispectrum fingerprint in the sense that an expansion of the LAD with respect to a basis set is sought; but the two methods differ in the use of radial basis functions in the Zernike case versus use of 4D spherical harmonics for the

bispectrum approach. This procedure leads to Zernike moments, c_{nl}^{m}, which are the coefficients in the expansion given by Equation 7.8 where Z_{nl}^{m} are Zernike basis functions. The coefficients are calculated using Equation 7.7. Invariance under rotation is then achieved by building the Zernike fingerprint with values corresponding to the norm of the vector $\boldsymbol{c}_{nl} = (c_{nl}^{-l}, c_{nl}^{-l+1}, ..., c_{nl}^{l})$ for different values of n and l.

$$c_{nl}^{m} = \langle Z_{nl}^{m}(\tilde{r}, \theta, \phi), \rho(\tilde{r}, \theta, \phi) \rangle \tag{7.7}$$

$$\rho(\tilde{r}, \theta, \phi) = \sum_{n=0}^{\inf} \sum_{l} \sum_{m=-l}^{l} c_{nl}^{m} Z_{nl}^{m}(\tilde{r}, \theta, \phi) \quad \text{for } \; n - l \geq 0 \tag{7.8}$$

Since the procedure is so similar to the one used for bispectrum fingerprints, Zernike fingerprints share the advantages offered by bispectrum fingerprints and are, moreover, more computationally efficient to calculate.

7.4.4 SOAP Fingerprints

The Smooth Overlap of Atomic Positions (SOAP) proposed by Bartók *et al.* [152] is a completely different approach to fingerprinting. In this method, a similarity measure (called SOAP) between two atomic environments given by their respective values of LAD ρ and ρ' is used directly for learning and predictions, instead of a descriptor. The SOAP, given by S, is defined as the inner product between the two LADs of the reference atoms (see Equation 7.9). The similarity kernel, k, is then obtained by integrating S over all possible rotations ($\hat{R}$) of one of the environments, as shown in Equation 7.10. A Gaussian-based LAD

given by Equation 7.11 is used for SOAP instead of Dirac delta function to avoid underestimating similarity between two slightly different environments. If the Gaussian-based LAD is expressed in terms of radial basis functions and spherical harmonics, it can be shown that the kernel given by Equation 7.10 becomes Equation 7.12, where b and b' are the bispectrum of the two environments as defined in Equation 7.6. This significant result shows that the need to construct descriptors can sometimes be circumvented with no loss of generality in favor of similarity measures that allow for direct comparison between atomic environments.

$$S(\rho, \rho') = \int \rho(r)\rho'(r)dr \tag{7.9}$$

$$k(\rho, \rho') = \int |S(\rho, \hat{R}\rho')|^n d\hat{R} \tag{7.10}$$

$$\rho(r) = \sum_i exp(-\alpha|r - r_i|^2) \tag{7.11}$$

$$k(\rho, \rho') = \sum_{n_1,n_2,n,l_1,l_2,l} b_{n_1 n_2 n l l_1 l_2} b'_{n_1 n_2 n l l_1 l_2} \tag{7.12}$$

7.5 Implementation of an open-source package for fingerprint computation

Given the increasing number of fingerprints proposed in the literature, it would clearly be desirable for potential users to have a software package that allows

quick and efficient evaluations of the suitability of different fingerprinting options for a given problem. Ideally, such a software package would include options to use all existing fingerprints. It should also be modular so that newly proposed fingerprints can be easily incorporated. Due to the computational cost of calculating some fingerprints, such as Bispectrums, efficient speed of computation would also be a requirement. Moreover, such a package should be easy to use and incorporate into the overall flow of using machine learning for molecular systems.

SEING is the name of a package that we have developed and are hereby releasing with those requirements in mind. SEING is written in C/C++ for fast computation of fingerprints and is designed in a modular fashion for extensibility. Packages such as AMP[134] and AeNET[174] and tensormol [175] include utilities for fingerprint calculations. But their primary focus is to use Neural Network approaches for machine learning force field development, whereas SEING focuses solely on the fingerprinting methods. As such, SEING allows more flexibility for the choice of which machine learning algorithm to use and allow applications beyond machine learning force field development.

An overview flowchart showing how SEING works is given in Fig. 7.3. The atomic coordinates of the system are typically provided in an XYZ file which is read and manipulated as an "AtomicSystem" object within SEING and used to instantiate the fingerprint calculator of interest. Support for other coordinate file formats will be added in the future. Other inputs such as the parameter values for the fingerprint of interest are provided in an input file. SEING also implements its own neighbor-searching algorithm for faster computation of local fingerprints. In SEING, every fingerprint is implemented as a separate calcula-

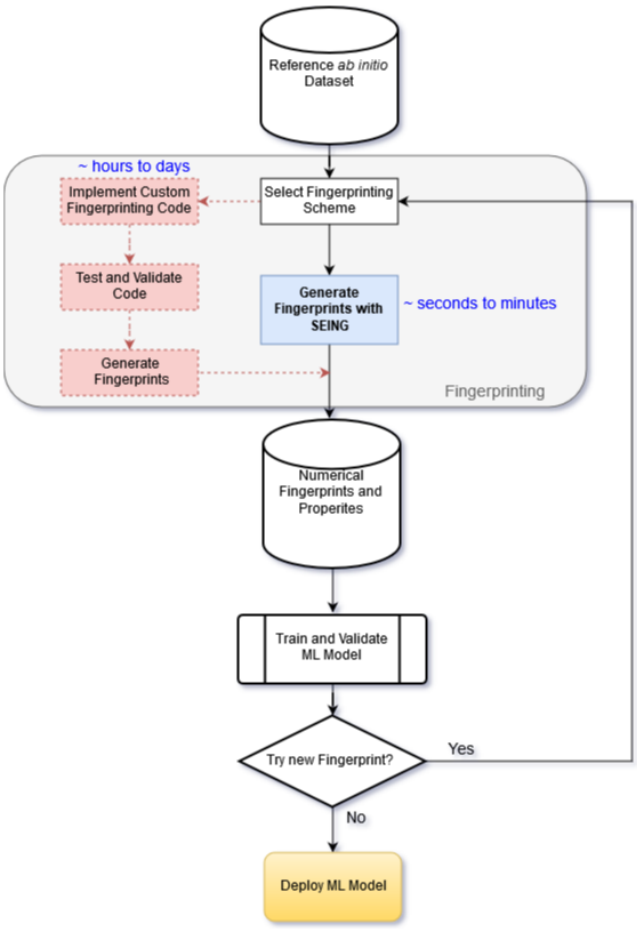


Figure 7.2: Comparison between the current way of conducting machine learning-based molecular studies (given in dashed lines and highlighted in red) and the proposed method using SEING (highlighted in blue). By providing a unified framework and package for fingerprint generation in an "off-the-shelf" fashion, SEING will make such machine learning studies more straightforward and effective.

tor. When a local fingerprint is needed, the "calculate_fingerprint" function of the calculator instance is called, with the atom of interest and its neighbors as arguments. In the case of a global fingerprint, the entire "AtomicSystem" object is used. From a development perspective, this allows any fingerprint-specific logic to be implemented within the calculator class which remains valid as long as an appropriate "calculate_fingerprint" function is exposed.

From a user's perspective, SEING has minimal requirements for installation and can be easily compiled on most operating systems. Using the code requires a coordinate file and an input file containing the type of fingerprint needed and any fingerprint-specific parameters. SEING implements two strategies to account for systems with multiple species: augmented and weighted. The "augmented" strategy increases the dimensionality of a given fingerprint by appending sub-fingerprints for each species and species combination whereas in the "weighted" strategy, any summation over atoms is modified by assigning a species-specific weight to each term. This weight can be the atomic number, electro-negativity, or any other value chosen by the user. Also, when available, derivatives of a fingerprint can be easily calculated and appended to the feature vector. More details on code installation, instructions for using the code, and how to contribute to the code are provided in the official documentation and user guide, accessible at https://seing.readthedocs.io. The source code is hosted on Github at https://github.com/mreveil/seing.

Our intention is that the availability of SEING will allow researchers to forgo a custom implementation of every fingerprint that they wish to use; this will allow them to focus on the predictive task at hand. Current fingerprinting methods implemented in SEING include Symmetry Functions, Bispectrum, AGNI

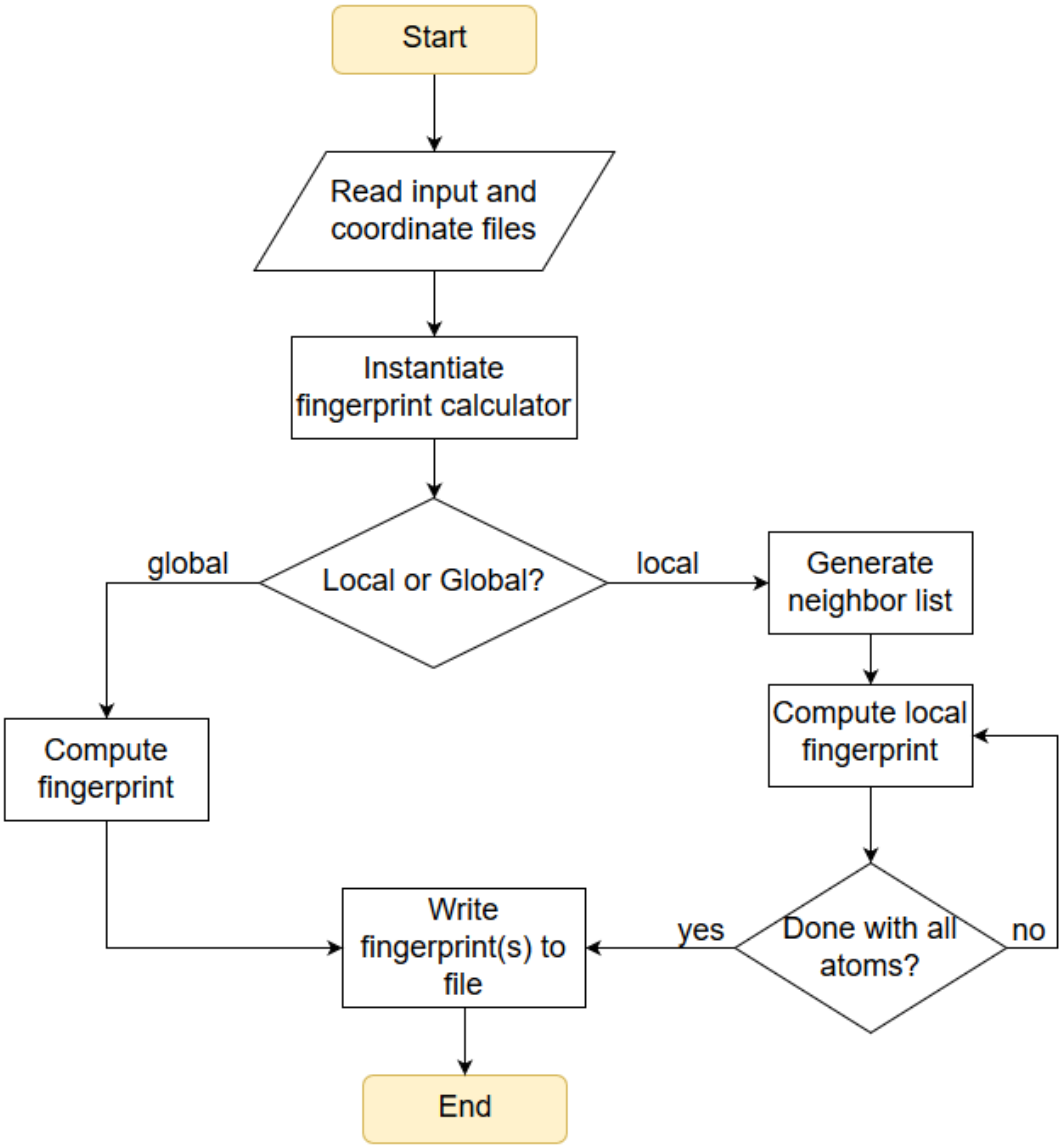


Figure 7.3: Flowchart showing the general procedure used by SEING to compute fingerprints.

and Zernike with more options in the pipeline for future additions. Since it is open-source, SEING also welcomes contributions from the community for bug-tracking and bug-solving, as well as implementation of new fingerprints and the addition of new features. In the next section, we will present examples of using SEING in a machine learning workflow.

CHAPTER 8

MACHINE-LEARNING DETERMINATION OF DIFFUSION PATHWAYS

8.1 Introduction

Two primary methods are available to conduct computational diffusion studies with atomistic-level resolution: Molecular Dynamics (MD) and Kinetic Monte Carlo (KMC). Given an initial configuration, in MD, Newton's equation of motion is solved for all the particles simultaneously to determine how their positions and velocities evolve with time. The mean square displacement with time and diffusivity can subsequently be calculated from the atomistic trajectory thus generated. One main limitation in using MD for diffusion studies is the limited timescale that is accessible to this method. Classical MD simulations are typically performed on the timescale of nanoseconds. In the case of *ab initio* MD, a more typical timescale is reduced to picoseconds as the force calculations become far more expensive. Consequently, for slow diffusing species, such as the ones of interest in this work, MD is not an effective tool to use as no significant particle motion will be observed during the duration of the simulation.

Kinetic Monte Carlo (KMC) is an attractive alternative to MD that allows access to much longer timescales and can therefore be effectively used to study slow diffusion processes. Kinetic Monte Carlo (KMC) is a rejection-free Monte Carlo method in which moves are chosen based on random sampling of the predefined probability distribution of the rates associated with all possible transitions in the system according to Eq. 8.1 where r_1 is a random number in the interval $[0, 1)$, m is the index of the chosen transition, k the transition rate and i, j, q are summation indices over all N possible events. Once an event/transition

is chosen, the molecular system is modified accordingly and the time is incremented by a value δt according to Eq. 8.2 where r_2 is a second random number in the interval (0,1).

$$\sum_{i=1}^{m-1} k_i \leq r_1 \sum_{j=1}^{N} k_j < \sum_{q=1}^{m} r_k \tag{8.1}$$

$$\delta t = -\frac{ln(r2)}{\sum_{j=1}^{N} r_j} \tag{8.2}$$

Kinetic Monte Carlo simulations are extremely successful in simulating a wide range of solid state diffusion and reaction problems. They give access to very long timescales on the order of seconds to minutes (and even hours). Such timescales are usually not accessible with other techniques such as Molecular Dynamics. The drawback with KMC, however, is the requirement to have rate information for all possible transitions in the system and the restriction (in the vast majority of implementations) of "jumps" to occur from one pre-defined lattice site to another. In most cases, it is difficult -if not impossible- to know in advance all the possible transitions that might take place, as well as their respective rates.

One common way to build such a rate library is to use DFT calculations. For example, in this work, we have used *ab initio* NEB calculations to calculate the energy barrier and associated Arrhenius-like rates for a number of predefined transitions. However, *ab initio* NEB calculations are computationally expensive. A typical energy barrier can cost around 2,500 CPU/hours to calculate for a 64-atom supercell using an Intel microprocessor and the Quantum Espresso software. Therefore, generating a large enough rate library for a complex system

like InGaAs can be prohibitively costly.

Here we develop a novel Machine Learning-based method which can be used to compute rates faster and cheaper compared to the traditional method. The goal is to use advanced machine learning techniques to predict the forces on each atom and couple this new capability with the NEB algorithm to predict the minimum energy path between two configurations.

8.2 Method description

Figure 8.1 shows a flowchart describing the proposed scheme. The process starts with data generation. In this first step, it is critical to sample sufficient points in a manner that is representative of the configuration space of the system of interest. *Ab initio* MD simulations are preferable to SCF calculations because they allow the system to explore more configurations around the starting configuration. This step leads to a collection of atomic configurations and associated forces and energy values that will be used as input for the machine learning process.

The second step is the feature engineering described in detail in Chapter 7. Several fingerprinting methods are available to transform the spatial coordinates into a format more amenable for use in machine learning. Those methods were reviewed and classified in Chapter 7 where a new custom-built software package called SEING was also presented as a means to streamline the process of computing fingerprints. It is important to note that inputs to a Machine Learning algorithm are generally referred to as "feature vectors" and each element of a given vector is a "feature."

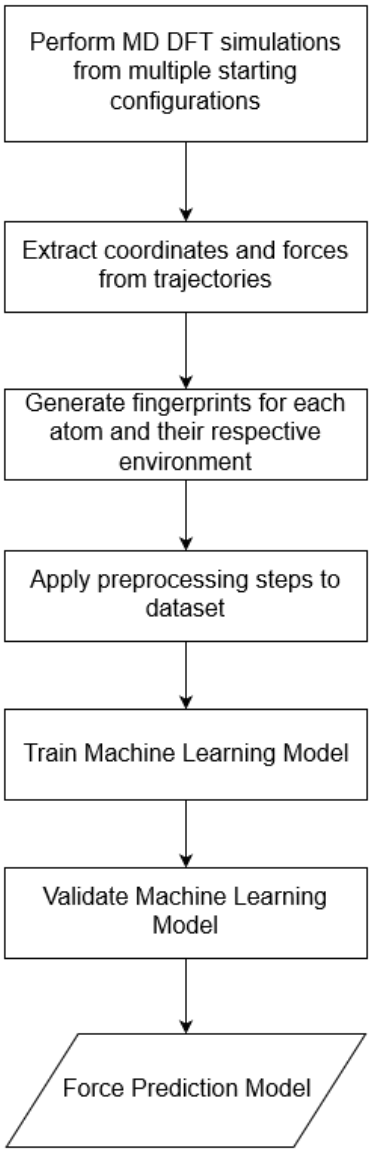


Figure 8.1: Flowchart showing the general procedure used to build a Machine Learning model to predict atomic forces on each atom.

Once the fingerprints are generated from cartesian coordinates, additional preprocessing steps are usually necessary to make the dataset ready for learning. For example, some algorithms such as artificial Neural Network perform better when all inputs are scaled to the same range. This avoids some features to misleadingly dominate over others and skew the learning process. In the case of multicomponent systems, it is also necessary to decide upon a method for incorporating the species for which the forces are being calculated as a feature.

After preparing the data in this manner, a suitable machine learning algorithm has to be used to build the predictive model. In Section 2.4, we presented a general background on machine learning methods and algorithms.

ML algorithms usually involve parameters that need to be tuned to the specific problem at hand. Since those parameters are different from the actual weights optimized during the fitting process, they are called hyperparameters and are usually optimized separately. For example, these hyper parameters might be the number of trees for a random forest approach, or the number of layers and neurons per layer for an artificial Neural Network.

Once a machine learning algorithm is selected and the hyperparameters optimized, the training process is performed on a subset of the data and a generalization score is calculated on a retained test subset of the data. In this work, we use Neural Network models and the R^2 score and Mean Absolute Error (MAE) to evaluate the ability of the Neural Networks developed to make predictions on future unseen data.

The final and key stage in this new method is using the model to make predictions of forces acting on a given atom in a certain configuration. This ML

force predictor is coupled with the NEB algorithm as shown in Fig. 8.1 to determine the minimum energy path between an initial and final configurations. Below we describe the implementation of this method for force prediction in InGaAs.

In summary, this method involves the following steps:

- DFT data generation: *ab initio* MD simulations are performed, followed by extraction of coordinates and associated forces from each frame. Details of those DFT MD simulations can be found in Section 8.3.
- Feature engineering: the SEING package is used to transform the Cartesian coordinates into fingerprints suitable for a machine learning application. Parameters for the fingerprint calculations are given in Table 8.1. Details on how to use SEING are available in the official documentation, accessible at seing.readthedocs.io
- Preprocessing: in this step, the data is scaled and split into training and validation sets. Details on the scaling and splitting are given in the last paragraph of Section 8.3 and in Section 8.4.1.
- Training: multiple machine learning algorithms are explored and the best is chosen, in this case Neural Network. The Neural Network optimization process is described in section 8.4.1.
- Validation: the best model is selected using the validation set. The same scaling used during training has to be applied to the validation set. The atomic forces are then computed using the ML model and compared to the expensive DFT results (See Section 8.4.2).
- Production: The by now trained model is used to make predictions during NEB simulations (See Section 8.4.3).

8.3 Simulation details

We perform DFT MD simulations with 30 different starting configurations of a 64-atom InGaAs supercell both with and without defects. To ensure the configuration space explored by each simulation is more extensive, an arbitrary temperature of 750 K is used. The starting configurations included defects such as interstitials and vacancies with some corresponding to saddle points identified from previous NEB calculations. Quantum Espresso was used for all the MD runs with Verlet integration and a Berendsen thermostat. Similarly to the NEB calculations, the DFT-MD runs were performed at the LDA level of theory.

Fingerprints are calculated using our custom-built SEING package which was described in Section 7.5. The Behler-Parinelo, zernike and AGNI fingerprints are used. Background details about those methods were given in Section 7.4. Table 8.1 gives parameter values used for each fingerprint. For the sake of simplification, as we only seek to provide a proof-of-concept for this method, silicon atoms are not included. No simulation was performed with Si.

Three distinct fingerprint computation strategies were used to account for the three species in InGaAs: augmented, weighted with atomic numbers and weighted with electronegativity. In the ”augmented” strategy, a separate sub-fingerprint was computed for each atom type or combination of atom types. For example, let’s assume the center atom is In. For the radial components, three sub-fingerprints are computed, one for each pair of atoms (In-In, In-Ga, In-As). For the angular components, six are calculated (In-In-In, In-In-Ga, In-In-As, Ga-In-Ga, Ga-In-As, As-In-As). All the sub-fingerprints are then concatenated to form the ”augmented” fingerprint. On the other hand, in the ”weighted” strat-

Fingerpint	Parameter	Value
Behler-Parinello	cutoff	6.5 Å
	η	0.05, 20.0, 50.0, 100.0
	γ	1.0, -1.0
	ζ	1.0, 4.0
	η_2	0.005
	append derivatives	yes
Zernike	cutoff	6.5 Å
	j_{max}	5
	n_{max}	5
	append derivatives	yes
AGNI	cutoff	6.5 Å
	$width$	0.1
	k_{max}	32
	append derivatives	no

Table 8.1: Parameter values used to compute the fingerprints for machine learning force prediciton in InGaAs. For more details on those fingerprints and the meaning of the parameters see Section 7.4.

egy, the contribution of each atom in all summations is assigned a weight. Two different types of weights are explore: the atomic number and the electronegativity of the atom.

Since artificial Neural Networks (which are used in this study) do not support categorical features, the atom type of the central atom must be converted into a continuous representation. A one-hot encoding method can be used to transform the three possible values (As, Ga or In) for the type of the central atom into a continuous representation that is then pre-pended to get final fingerprint. One-hot encoding works by creating a number of dummy binary variables equal to the number of categories. Then, for each category, a different binary variable is set to 1 while the others are set to zero. This method of assigning a numerical value to the categorical features is advantageous because it does not assume any ordering between the different categories. In Table 8.2, we show the complete list of all the datasets obtained in this manner.

A total of about 110,000 data points are extracted from the molecular dynamics simulation followed by fingerprint computation. Based on the stoichiometry of InGaAs, half of those data points are for As atoms and the other half split evenly between Ga and In. A total of at least 40,000 data points (30% of the total dataset), split between As (50%), Ga (25%) and In (25%), are randomly selected and set apart for testing. The remaining data points are left for training purposes. The test set is used to calculate the generalization score after a final trained model is obtained.

The neuralnet [176] package implemented in the statistical framework R [177] is used for the Neural Network training. The column-wise spread and means of the training dataset are calculated and used to scale the data (see

Model #	Fingerprint	Strategy	Size
1	Behler-Parinello	augmented	144
2	Behler-Parinello	weighted (atomic number)	64
3	Behler-Parinello	weighted (electronegativity)	64
4	Zernike	weighted (atomic number)	72
5	Zernike	weighted (electronegativity)	72
6	AGNI	augmented	96
7	AGNI	weighted (atomic number)	64
8	AGNI	weighted (electronegativity)	64

Table 8.2: Fingerprint size for each of the eight types of machine learning models trained in this work. To speed up the process, a parallel processing approach is adopted whereby a specific model is trained for each atom type (In, Ga or As) and each force direction (x, y or z). This lead to a total of 72 different models.

Equation 8.3 where X is the feature vector, M and S are, respectively, vectors of means and spread for each feature) and then stored. At prediction time, each data point is similarly scaled using the spread and means calculated from the training dataset. This process leads to features centered around zero.

$$X_{scaled} = \frac{X - M}{S} \tag{8.3}$$

8.4 Results and discussion

8.4.1 Effect of dataset size and network architecture

We explore the effects of the size of the training dataset on the generalization error. This step ensures enough data points are included in the training process. We vary the number of data points included in the training set from 1,000 to 15,000 in steps of 1,000. The training subset is selected by random sampling from the 30,000 points set apart for training. Throughout this process we do not include any datapoint from the 20,000 data points of the test subset. A Neural Network with three layers of 10 neurons each is used and the training is performed until the derivative of the gradient reaches a threshold of 0.01. The Behler-Parinello fingerprint is used for this set of study.

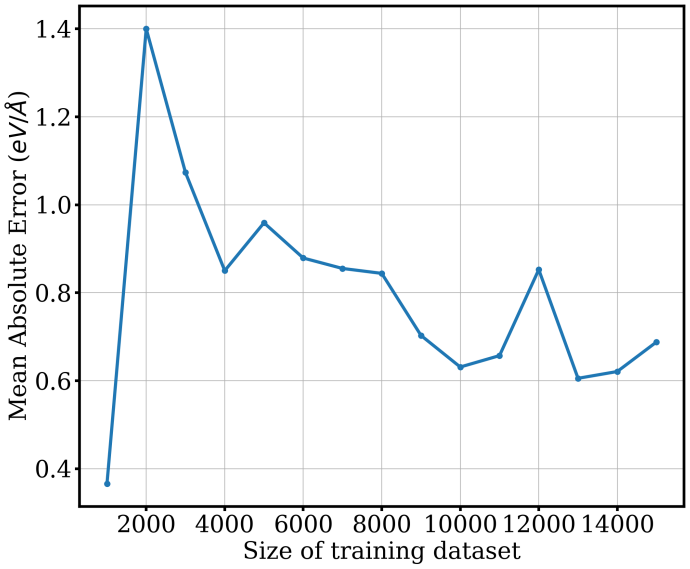
After the training process is completed, the resulting model is used to predict the forces for the 20,000 data points in the test subset. The results are summarized in Fig. 8.2 and show that the R^2 error quickly increases with system size until it reaches a value of 0.81 for a training dataset size of 10,000 points. Sim-

ilarly, the mean absolute error quickly decreases with system size until a value of 0.62 is reached for a dataset of 10,000 points. Only marginal improvement is observed for both the R^2 and MAE scores beyond the 10,000 data points mark. This means that 10,000 data points are sufficient to train the Neural Network.

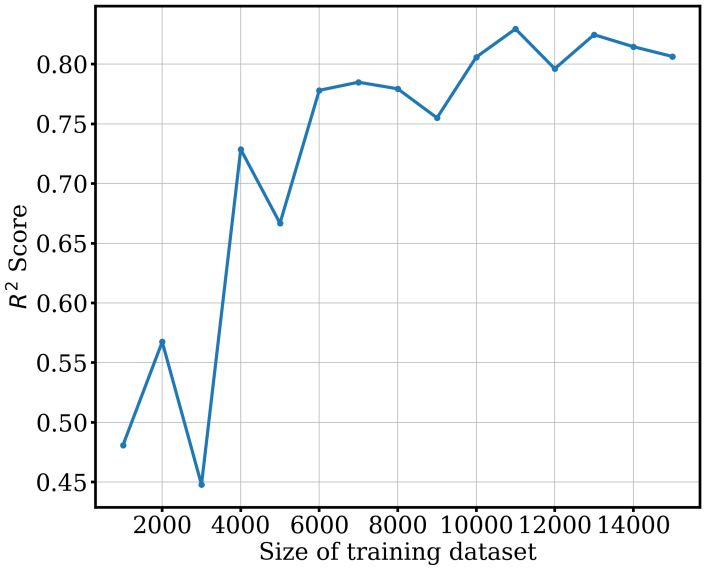
To find the optimal ANN architecture as defined by the number of layers and neurons per layer, we vary the number of layers from 1 to 12 in increments of 1, and the number of neurons from 10 to 30 in increments of 5. We use 10,000 data points, randomly selected from the training subset, to train the Neural Networks. Our results (see Fig. 8.3) show that a network with three layers of 25 neurons each leads to the best results in terms of the R^2 score, and the MAE and RMS scores achieved.

8.4.2 Machine learning models for force prediction

We use 20,000 and 10,000 data points, respectively, to train Neural Networks for As and In/Ga. Each Neural Network has three layers with 25 neurons each, as per the optimization approach reported in the previous section. Separate models are created for each atom for faster computation. We also train models for each component of the forces separately. This leads to a total of nine Neural Network models (three atoms with three force components each), which are then tested on 20,000 and 10,000 data points, respectively, for As and Ga/In. The results are plotted in Fig. 8.4 and show excellent predictive capabilities as given by the R^2, MAE and RMS scores of our predictions. On average, an MAE error of 0.08 eV/Å is obtained, comparable to the expected error of DFT calculations.

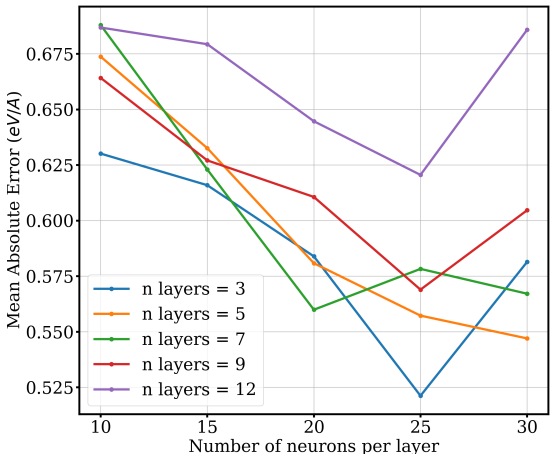


(a)

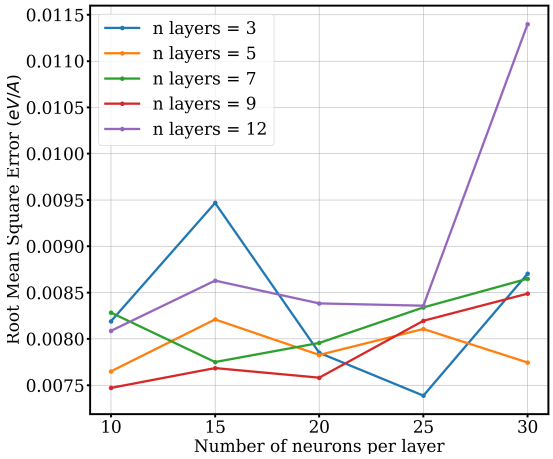


(b)

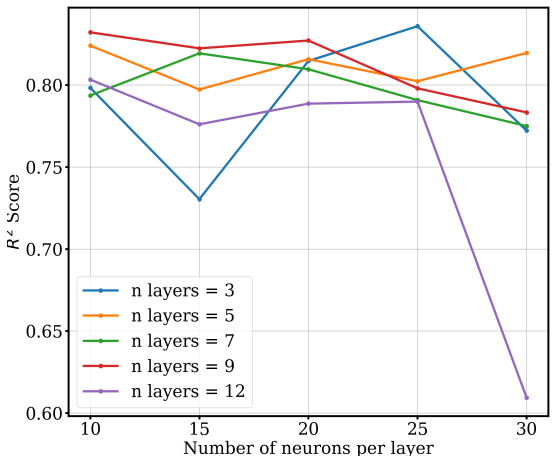
Figure 8.2: Mean Absolute Error (MAE) (a) and R^2 scores (b) of a Neural Network model for force prediction on As atoms as a function of dataset size. The scores are calculated based on a held-out test set of 20,000 data points. The results show that 10,000 points or more are needed to train the Neural Network for proper generalization.



(a)



(b)



(c)

Figure 8.3: Mean Absolute Error (a), Root Mean Square Error (b) and R^2 scores (c) as a function of number of neurons for force prediction on a retained set of 20,000 data points. Each line represents a network with a different number of layers.

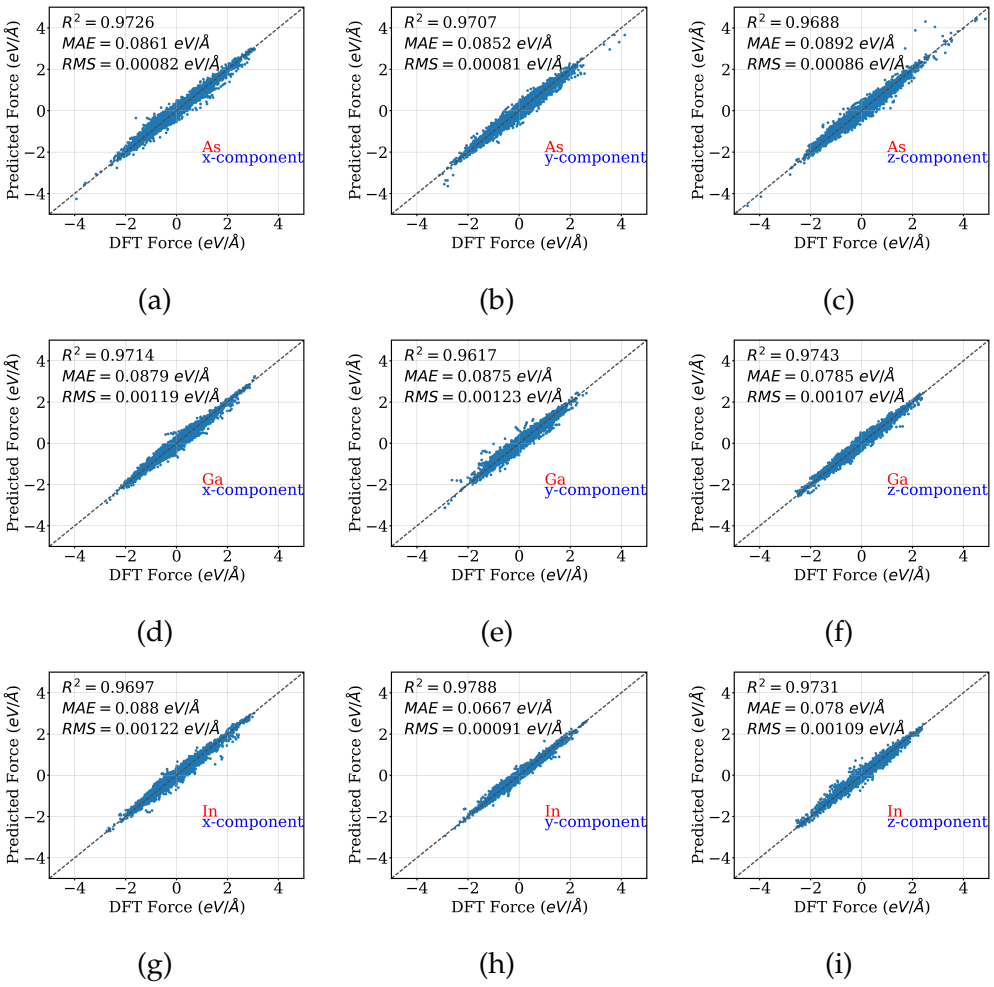


Figure 8.4: Performance of Neural Network machine learning models for force prediction in InGaAs crystals with Behler-Parinelo fingerprints computed with SEING, our custom-made software package for fingerprint computing.

8.4.3 Predictions of diffusion pathways

We then use the trained Neural Network models in a new framework that we call "Machine Learning NEB" or "ML NEB" to compute the diffusion pathways for several point defects in InGaAs. The NEB code published by Herbol *et al.* [60] was used and adapted for this purpose. We start by linearly interpolating between the initial and final configurations of a given atomistic jump to create guessed intermediate states for the diffusion pathway. We then computed fingerprints for each atom in each intermediate image using SEING [132]. We then fed those fingerprints to the Neural Network models previously created to predict forces on each atom. Finally, we use these Machine Learning predicted forces within the framework of the NEB algorithm to nudge the diffusion pathway towards a more favorable pathway. We then go back to computing the fingerprints for the new atomic positions and repeat the steps until convergence. This process is illustrated in Fig. 8.5.

Table 8.3 lists the different transitions used to test our new method. We include split interstitials as well as second nearest-neighbor vacancy-assisted jumps of Ga, As and In. For each of those transitions, after the ML NEB computations converge, we calculate the distance between each atom in the ML pathway and their counterpart in the DFT pathway. Those distances are then averaged over the number of atoms and states to provide the overall average deviation from DFT NEB trajectories. Similarly, we compute the average deviation of the moving atom's trajectory by considering only the distances between the moving atom in the ML and DFT NEB pathways.

Values obtained for both deviations are plotted in Fig. 8.6 for all ten transitions. In all cases, the overall average deviation of the ML NEB runs com-

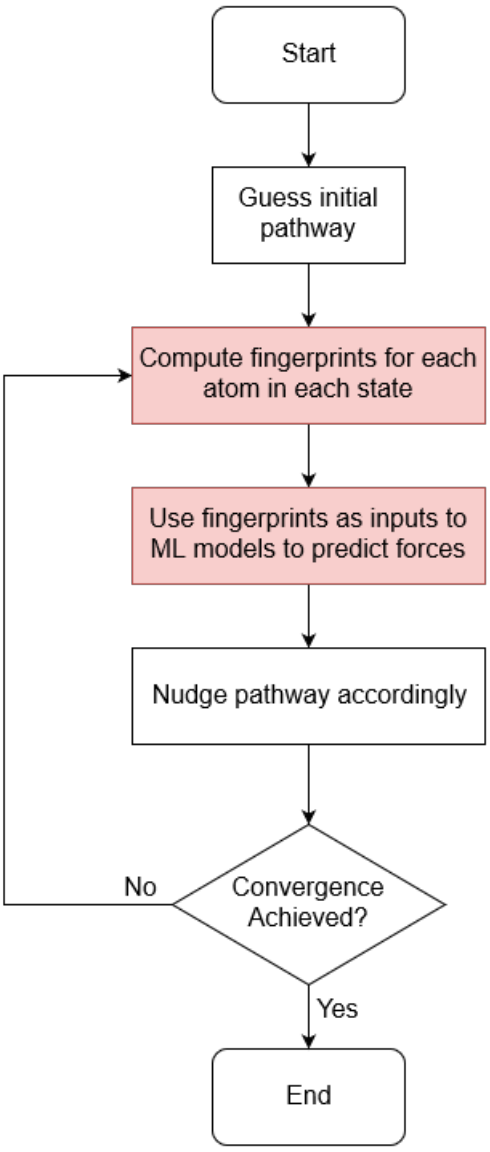


Figure 8.5: Flowchart describing the new Machine Learning NEB framework used in this work. This assumes that a suitable Machine Learning force prediction model has already been properly trained.

pared to DFT NEB computations is less than 0.2 Å. Most of this deviation can be attributed to the deviations of the moving atom which vary from 0.1 Å to 0.8 Å. Understandably, deviations from DFT NEB trajectories tend to be higher for transitions that involve a longer distance as there is more room for residual movement. Overall, deviations of less than 0.8 Å are more than acceptable, specially given the significant savings in computation time and resources afforded by this new method which yields the transition pathway within minutes on one core compared to tens of hours on 56 cores.

8.5 Limitations and future work

Energy values

This first iteration of our ML NEB method does not include energies. Moving forward, there are two options for improvement: One method is to train separate Machine Learning energy models and couple them within the current method to automatically calculate energies for each state. This method will introduce an additional source of uncertainty as there will be intrinsic errors associated with the ML energy models. Another method is to perform simple DFT calculations for each state in the optimized transition pathway to generate an energy profile along the reaction coordinates. Single computations of energies with DFT would be relatively cheap and would come with the advantage of added accuracy.

Transition	Mechanism	Moving Atom	Initial Site	Final Site
S1	Vacancy-Assisted	Ga	Ga	In
S2	Vacancy-Assisted	As	As	As
S3	Vacancy-Assisted	As	Ga	As
S4	Vacancy-Assisted	As	In	As
S5	Split-interstitial	As	S_1	S_1
S6	Split-interstitial	As	S_2	S_2
S7	Split-interstitial	As	S_2	S_3
S8	Split-interstitial	As	S_3	S_3
S9	Vacancy-Assisted	In	Ga	Ga
S10	Vacancy-Assisted	In	Ga	In

Table 8.3: List of the transitions selected for comparison between DFT and ML NEB diffusion results.

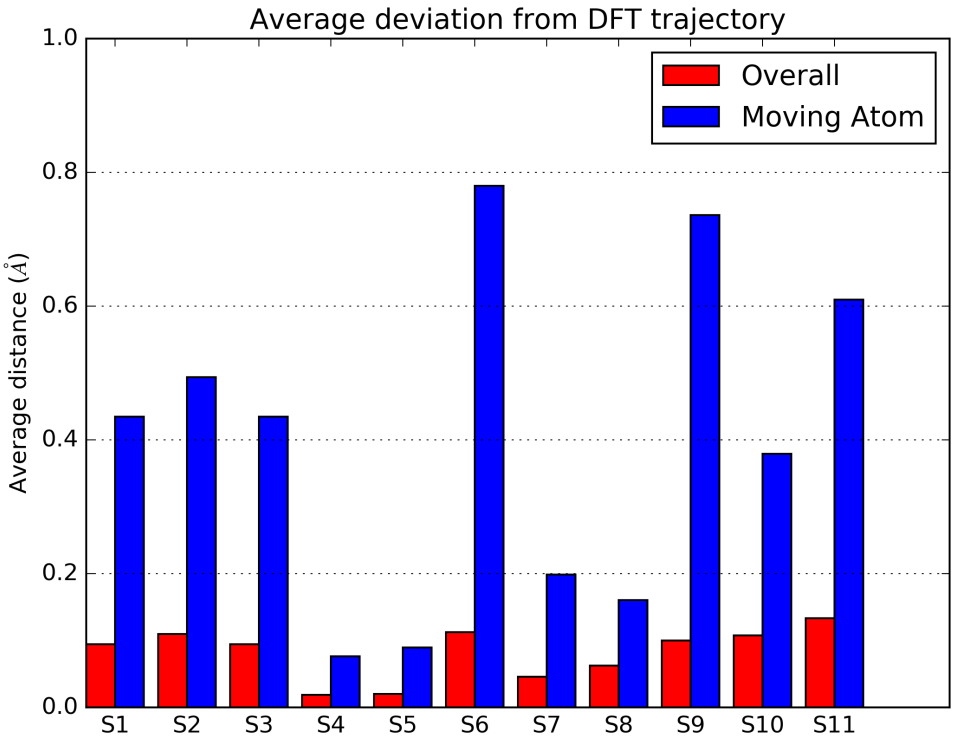


Figure 8.6: Comparison between the diffusion paths obtained through DFT calculations and our ML method expressed in terms of distance between the two paths (or deviations from the DFT path).

Geometry optimization

One limitation of the current iteration of our method is the use of DFT-generated geometries for the initial and final states. Although the cost for these geometry optimizations is essentially negligible compared to the NEB runs, it would be advantageous to be able to perform them using the capability to predict forces with the ML models. This would increase the time-savings at the cost of potentially less accuracy.

8.6 Conclusions

In this chapter, we presented a novel approach to study diffusion using advanced machine learning techniques. This new method involves computation of a dataset of atomic configurations and associated forces used to train Neural Networks for force predictions on new configurations. These Machine Learning models are then coupled with NEB to find the minimum energy path for a given transition. We apply this method to point-defect diffusion in InGaAs and compare our results to previous DFT calculations for a total of ten transitions. Our results show that this method can predict diffusion pathways with reasonable accuracy at a fraction of the cost compared to DFT.

CHAPTER 9

LIMITATIONS AND FUTURE WORK

This work covers intrinsic and silicon dopant diffusion in three separate III-V semiconductors and develops an improved method to calculate diffusion rates that can save significant computing power while maintaining reasonable accuracy. However, there are a few limitations and some suggestions for related future work that are worth noting.

9.1 Experimental validation

The energy barriers computed and the diffusion models developed in this work are mostly theoretical. Experimental validation would undoubtedly increase the value and applicability of those new insights gained from DFT calculations. In particular, it would be great to have experimental verification of the presence of split interstitial species which has been reported in InGaAs for the first time in our studies and previously in GaAs by other authors. In addition, *in situ* diffusion experiments that could confirm the proposed role of Si pairs and split interstitial species in the enhanced diffusion at high concentration would provide strong supporting evidence for our model. Also, most of the diffusion experiments reported for the In, Ga and As family of III-Vs in literature were determined 40 years ago. The expertise acquired by the community, as well as the technological advancements in laboratory equipment and methods with atomic resolution, could help conduct more accurate and comprehensive experiments to fully explore and characterize the concentration dependence of Si diffusion in III-V materials.

9.2 Modeling of random InGaAs crystals

As mentioned in Section 5.2.1, the $In_{0.5}Ga_{0.5}As$ crystal preferentially adopts a random ordering. The CuAu-I ordering studied in this work exists only under very specific conditions, such as MBE growth, as it is not the most thermodynamically favorable state [92]. Therefore, diffusion studies in the random crystal and comparisons with those in the CuAu-I ordering would be useful. There are, however, a couple of serious challenges to perform such simulations.

The first challenge is related to system size. As mentioned in Chapter 2, DFT calculations are expensive and usually limited to system sizes on the order of a few hundreds of atoms. For instance, due to the large number of simulations performed in this work, we were able to simulate only crystals containing 64 ± 1 atoms. Such a small system size would be too small for a random crystal. An even bigger system size (compared to ordered crystals) will be necessary if periodic boundary conditions are to be used with random crystals.

The second challenge is related to the calculation of enough energy barriers to build the rate library necessary to perform Kinetic Monte Carlo studies. Since the cations in the crystal are randomly situated, local environments around a diffusing species will be often different from previously seen local environments. This makes it practically impossible to pre-build a library of rates and use a method such as Kinetic Monte Carlo to study the diffusion of specific defects.

The Machine Learning method developed in this work could open the door to perform such studies. With the decrease in computing power and time required to calculate an energy barrier between two configurations, ML NEB runs

can be performed on the fly. The new algorithm would involve isolating local environments around the diffusing species, computing new energy barriers and updating the rate library at regular intervals and on-the-fly as the Kinetic Monte Carlo simulation progresses in a self-learning fashion.

9.3 Beyond the In-Ga-As family of III-V materials

We believe our work with the In-Ga-As family of III-V semiconductors, in addition to the new insights we have gained, have laid the ground work for the study of more complex alloys such as quaternaries and so on. Beyond the band gap engineering opportunity offered by those more complex III-V materials, they may also allow interesting diffusion and activation processes to take place. The interplay between intrinsic defects, dopant atoms and other minority atoms in compositions with more than three species could perhaps hold the key to overcoming the limitation created by the compensation effects in GaAs and InGaAs. We believe our continuum model would provide a suitable theoretical framework for such studies while our machine learning approach could help accelerate the exploration of different diffusion pathways to find the most favorable ones.

CHAPTER 10

CONCLUSIONS

In this thesis, we performed a large number of NEB calculations at the DFT level to explore different diffusion mechanism of intrinsic and Si defects in III-V materials. These studies were primarily meant to uncover the atomistic origin of the concentration-dependence of Si diffusion in GaAs and InGaAs.

Our results provide additional insights that go beyond the early assumptions of concentration-dependence diffusivity driven by Ga vacancies. Instead, we show that the enhanced diffusion is due to highly mobile Si species that appear at high concentration. We show that Si-Si complexes and Si split interstitials that would form at high Si concentration have higher diffusivity and could explain the overall increase in diffusion coefficient. We use these results to build an analytical model to study Si diffusion in III-Vs at the continuum level. We validate the model using comparisons with published experimental data. We also explain how those new species that exist at high Si concentration could lead to the observed saturation of carrier concentration.

Moreover, our progress in incorporating advanced Machine Learning capabilities in molecular simulation were reported in this thesis. We show that this new ML NEB approach can save significant time and resources while yielding reasonable accuracy for the determination of diffusion pathways. Such a method could be used as a screening technique to quickly gain an understanding of the relative importance of different diffusion mechanisms. It could also be coupled with KMC simulations in a self-learning fashion as the speed at which the ML NEB method allows diffusion rates to be extracted affords on-the-fly calculations. We hope this work on diffusion in the In-Ga-As family of III-Vs

has not only increased understanding of diffusion in this class of materials but also laid the ground work for the study of more complex III-V compounds.

BIBLIOGRAPHY

[1] H. L. Aldridge Jr., A. G. Lind, M. E. Law, C. Hatem, and K. S. Jones, "Concentration-dependent diffusion of ion-implanted silicon in $In_{0.53}Ga_{0.47}As$," *Appl. Phys. Lett.*, vol. 105, p. 042113, 2014.

[2] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electronics*, vol. 20, no. 2, pp. 77 – 89, 1977.

[3] R. N. Hall, "Electron-hole recombination in germanium," *Phys. Rev.*, vol. 87, pp. 387–387, Jul 1952.

[4] S. Mokkapati and C. Jagadish, "Iii-v compound sc for optoelectronic devices," *Materials Today*, vol. 12, no. 4, pp. 22 – 32, 2009.

[5] C. C. Stoumpos and M. G. Kanatzidis, "Halide perovskites: Poor man's high-performance semiconductors," *Advanced Materials*, vol. 28, no. 28, pp. 5778–5793, 2016.

[6] Z. Xiao, Y. Zhou, H. Hosono, T. Kamiya, and N. P. Padture, "Bandgap optimization of perovskite semiconductors for photovoltaic applications," *Chemistry A European Journal*, vol. 24, no. 10, pp. 2305–2316, 2018.

[7] E. Fortunato, P. Barquinha, and R. Martins, "Oxide semiconductor thin-film transistors: A review of recent advances," *Advanced Materials*, vol. 24, no. 22, pp. 2945–2986, 2012.

[8] R. R. Schaller, "Moore's law: past, present and future," *IEEE Spectrum*, vol. 34, pp. 52–59, June 1997.

[9] K. Lee, I. Nam, I. Kwon, J. Gil, K. Han, S. Park, and B.-I. Seo, "The impact of semiconductor technology scaling on cmos rf and digital circuits for wireless application," *IEEE Transactions on Electron Devices*, vol. 52, pp. 1415–1422, July 2005.

[10] M. Green, "Solar cells: operating principles, technology, and system applications,"

[11] G. Niu, X. Guo, and L. Wang, "Review of recent progress in chemical stability of perovskite solar cells," *J. Mater. Chem. A*, vol. 3, pp. 8970–8980, 2015.

[12] M. A. Green, "Thin-film solar cells: review of materials, technologies and commercial status," *Journal of Materials Science: Materials in Electronics*, vol. 18, pp. 15–19, Oct 2007.

[13] I. D. W. Samuel and G. A. Turnbull, "Organic semiconductor lasers," *Chemical Reviews*, vol. 107, no. 4, pp. 1272–1295, 2007. PMID: 17385928.

[14] Z. Yin and X. Tang, "A review of energy bandgap engineering in iiiv semiconductor alloys for mid-infrared laser applications," *Solid-State Electronics*, vol. 51, no. 1, pp. 6 – 15, 2007.

[15] G. H. M. van Tartwijk and D. Lenstra, "Semiconductor lasers with optical injection and feedback," *Quantum and Semiclassical Optics: Journal of the European Optical Society Part B*, vol. 7, no. 2, p. 87, 1995.

[16] D. Delbeke, R. Bockstaele, P. Bienstman, R. Baets, and H. Benisty, "High-efficiency semiconductor resonant-cavity light-emitting diodes: a review," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 8, pp. 189–206, Mar 2002.

[17] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. John wiley & sons, 2006.

[18] D. K. Schroder, *Semiconductor material and device characterization*. John Wiley & Sons, 2006.

[19] M. D. McCluskey and E. E. Haller, *Dopants and defects in semiconductors*. CRC Press, 2018.

[20] T. F. Kuech, "Iii-v compound semiconductors: Growth and structures," *"Progress in Crystal Growth and Characterization of Materials*, vol. 62, no. 2, pp. 352 – 370, 2016. Special Issue: Recent Progress on Fundamentals and Applications of Crystal Growth; Proceedings of the 16th International Summer School on Crystal Growth (ISSCG-16).

[21] X. Peng, J. Wickham, and A. Alivisatos, "Kinetics of ii-vi and iii-v colloidal semiconductor nanocrystal growth:focusing of size distributions," *Journal of the American Chemical Society*, vol. 120, no. 21, pp. 5343–5344, 1998.

[22] W. W. Yu and X. Peng, "Formation of high-quality cds and other ii–vi semiconductor nanocrystals in noncoordinating solvents: tunable reactiv-

ity of monomers," *Angewandte Chemie International Edition*, vol. 41, no. 13, pp. 2368–2371, 2002.

[23] L. I. Berger, *Semiconductor materials*. CRC press, 1996.

[24] A. Stein, G. A. Ozin, and G. D. Stucky, "From the molecule to an expanded i-vii semiconductor quantum superlattice: silver, sodium halo-sodalites," *Journal of the American Chemical Society*, vol. 112, no. 2, pp. 904–905, 1990.

[25] C.-Y. Yeh, Z. Lu, S. Froyen, and A. Zunger, "Zinc-blende–wurtzite polytypism in semiconductors," *Physical Review B*, vol. 46, no. 16, p. 10086, 1992.

[26] I. Vurgaftman, J. Meyer, and L. Ram-Mohan, "Band parameters for iii–v compound semiconductors and their alloys," *Journal of applied physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

[27] S. Oktyabrsky and D. Y. Peide, *Fundamentals of III-V semiconductor MOSFETs*. Springer, 2010.

[28] J. A. del Alamo, "Nanometre-scale electronics with III-V compound semiconductors," *Nature*, vol. 479, p. 317, 2011.

[29] S. Nakamura, "Iii–v nitride based light-emitting devices," *Solid State Communications*, vol. 102, no. 2-3, pp. 237–248, 1997.

[30] M. Yamaguchi, "Iii–v compound multi-junction solar cells: present and future," *Solar energy materials and solar cells*, vol. 75, no. 1-2, pp. 261–269, 2003.

[31] ITRS, "The international technology roadmap for semiconductors (ITRS)," 2011.

[32] A. Seeger and K. Chik, "Diffusion mechanisms and point defects in silicon and germanium," *physica status solidi (b)*, vol. 29, no. 2, pp. 455–542, 1968.

[33] B. Tuck, *Atomic diffusion in III-V semiconductors*. CRC Press, 1988.

[34] D. Larkin, "Sic dopant incorporation control using site-competition cvd," *physica status solidi (b)*, vol. 202, no. 1, pp. 305–320, 1997.

[35] J. F. Ziegler, *Ion implantation science and technology*. Elsevier, 2012.

[36] J. F. Gibbons, "Ion implantation in semiconductorspart ii: Damage production and annealing," *Proceedings of the IEEE*, vol. 60, no. 9, pp. 1062–1096, 1972.

[37] E. Dimakis, M. Ramsteiner, C.-N. Huang, A. Trampert, A. Davydok, A. Biermanns, U. Pietsch, H. Riechert, and L. Geelhaar, "In situ doping of catalyst-free InAs nanowires with Si: Growth, polytypism, and local vibrational modes of si," *Applied Physics Letters*, vol. 103, no. 14, p. 143121, 2013.

[38] D. W. Park, S. G. Jeon, C.-R. Lee, S. J. Lee, J. Y. Song, J. O. Kim, S. K. Noh, J.-Y. Leem, and J. S. Kim, "Structural and electrical properties of catalyst-free Si-doped InAs nanowires formed on Si(111)," *Scientific Reports*, vol. 5, 2015.

[39] K.-F. Wang, Y. Gu, X. Yang, T. Yang, and Z. Wang, "Si delta doping inside InAs/GaAs quantum dots with different doping densities," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 30, no. 4, p. 041808, 2012.

[40] R. C. Keller and C. R. Helms, "Oxidation enhanced diffusion of Si in GaAs: The effect of excess As on diffusion depth and carrier concentration," *Applied Physics Letters*, vol. 67, no. 3, pp. 398–400, 1995.

[41] S. Nakamura, T. Mukai, and M. Senoh, "Si- and ge-doped gan films grown with gan buffer layers," *Japanese Journal of Applied Physics*, vol. 31, no. 9R, p. 2883, 1992.

[42] S. Fritze, A. Dadgar, H. Witte, M. Bgler, A. Rohrbeck, J. Blsing, A. Hoffmann, and A. Krost, "High si and ge n-type doping of gan doping - limits and impact on stress," *Applied Physics Letters*, vol. 100, no. 12, p. 122104, 2012.

[43] D. C.J., L. C., C. M.G., S. L.M., R. S., W. A., and W. I.M., "Silicon doping of gallium nitride using ditertiarybutylsilane," *Chemical Vapor Deposition*, vol. 10, no. 4, pp. 187–190, 2004.

[44] V. Sorg, S. N. Zhang, M. Hill, P. Clancy, and M. O. Thompson, "(invited) dopant activation and deactivation in InGaAs during sub-millisecond thermal annealing," *ECS Transactions*, vol. 21, 2015.

[45] A. Lind, M. A. Gill, C. Hatem, and K. S. Jones, "Electrical activation of

ion implanted Si in amorphous and crystalline $In_{0.53}Ga_{0.47}As$," *Nucl. Instr. Meth. Phys. Res. B*, vol. 337, pp. 7–10, 2014.

[46] A. G. Lind, N. G. Rudawski, N. J. Vito, C. Hatem, M. C. Ridgway, R. Hengstebeck, B. R. Yates, , and K. S. Jones, "Maximizing electrical activation of ion-implanted si in $In_{0.53}Ga_{0.47}As$," *Appl. Phys. Lett.*, vol. 103, p. 232102, 2013.

[47] L. Vieland, "The effect of arsenic pressure on impurity diffusion in gallium arsenide," *Journal of Physics and Chemistry of Solids*, vol. 21, no. 3, pp. 318 – 320, 1961.

[48] G. Antell, "The diffusion of silicon in gallium arsenide," *Solid-State Electronics*, vol. 8, pp. 943–944, 1965.

[49] K. L. Kavanagh, J. W. Mayer, C. W. Magee, J. Sheets, J. Tong, and J. M. Woodall, "Silicon diffusion at polycrystalline-si/gaas interfaces.," *Applied Physics Letters*, vol. 47, no. 11, p. 1208, 1985.

[50] E. F. Schubert, J. B. Stark, T. H. Chiu, and B. Tell, "Diffusion of atomic silicon in gallium arsenide," *Applied Physics Letters*, vol. 53, no. 4, pp. 293–295, 1988.

[51] D. G. Deppe and N. H. Jr., "Atom diffusion and impurityinduced layer disordering in quantum well iiiv semiconductor heterostructures," *Journal of Applied Physics*, vol. 64, no. 12, pp. R93–R113, 1988.

[52] S. Yu, U. M. Gsele, and T. Y. Tan, "A model of si diffusion in gaas based on the effect of the fermi level," *Journal of Applied Physics*, vol. 66, no. 7, pp. 2952–2961, 1989.

[53] K. S. Jones, A. G. Lind, C. Hatem, S. Moffatt, and M. C. Ridgeway, "Brief review of doping issues in III-V semiconductors," *ECS Trans.*, vol. 53, p. 97, 2013.

[54] A. G. Lind, H. L. A. Jr., C. Hatem, M. E. Law, and K. S. Jones, "Review-dopant selection considerations and equilibrium thermal processing limits for n+-$In_{0.53}Ga_{0.47}As$," *ECS J. Solid State Sci. Technol.*, vol. 5, p. Q125, 2016.

[55] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.

[56] G. Henkelman, G. Jóhannesson, and H. Jónsson, *Methods for Finding Saddle Points and Minimum Energy Paths*, pp. 269–302. Springer Netherlands, 2002.

[57] G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," *J. Chem. Phys.*, vol. 113, p. 9901, 2000.

[58] H. Jónsson, G. Mills, and K. W. Jacobsen, *Nudged elastic band method for finding minimum energy paths of transitions*. Singapore: World Scientific, 1998.

[59] G. Mills and H. Jónsson, "Quantum and thermal effects in H_2 dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems," *Physical Review Letters*, vol. 72, 1994.

[60] H. C. Herbol, J. Stevenson, and P. Clancy, "Computational implementation of nudged elastic band, rigid rotation, and corresponding force optimization," *Journal of Chemical Theory and Computation*, vol. 13, no. 7, pp. 3250–3259, 2017.

[61] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 89 – 109, 2001.

[62] K. K. Wong, L. Wang, and D. Wang, "Recent developments in machine learning for medical imaging applications," *Computerized Medical Imaging and Graphics*, vol. 57, no. Supplement C, pp. 1 – 3, 2017. Recent Developments in Machine Learning for Medical Imaging Applications.

[63] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," *Computational and Structural Biotechnology Journal*, vol. 13, no. Supplement C, pp. 8 – 17, 2015.

[64] A. Graves, A. r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645–6649, May 2013.

[65] L. Deng, G. Hinton, and B. Kingsbury, "New types of deep neural network learning for speech recognition and related applications: an overview," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8599–8603, May 2013.

[66] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, and A. Y. Ng, "Deep speech: Scaling up end-to-end speech recognition," *ArXiv e-prints*, dec 2014.

[67] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," *ArXiv e-prints*, Nov. 2014.

[68] S.-J. Wang, H.-L. Chen, W.-J. Yan, Y.-H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine," *Neural Processing Letters*, vol. 39, pp. 25–43, Feb 2014.

[69] S. E. Thompson, F. Mullally, J. Coughlin, J. L. Christiansen, C. E. Henze, M. R. Haas, and C. J. Burke, "A machine learning technique to identify transit shaped signals," *The Astrophysical Journal*, vol. 812, no. 1, p. 46, 2015.

[70] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, and O. A. von Lilienfeld, "Prediction errors of molecular machine learning models lower than hybrid dft error," *Journal of Chemical Theory and Computation*, vol. 13, no. 11, pp. 5255–5264, 2017. PMID: 28926232.

[71] M. Fernandez, N. R. Trefiak, and T. K. Woo, "Atomic property weighted radial distribution functions descriptors of metalorganic frameworks for the prediction of gas uptake capacity," *The Journal of Physical Chemistry C*, vol. 117, no. 27, pp. 14095–14105, 2013.

[72] W. W. Tipton and R. G. Hennig, "A grand canonical genetic algorithm for the prediction of multi-component phase diagrams and testing of empirical potentials," *Journal of Physics: Condensed Matter*, vol. 25, no. 49, p. 495401, 2013.

[73] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, "Aflowlib.org: A distributed materials properties repository from high-throughput *ab initio* calculations," *Computational Materials Science*, vol. 58, no. Supplement C, pp. 227 – 235, 2012.

[74] B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rhl, and C. Wolverton, "The open quantum materials database (oqmd): assessing the accuracy of dft formation energies," *Npj Computational Materials*, 12 2015.

[75] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, "Finding natures missing ternary oxide compounds using machine learning and density functional theory," *Chemistry of Materials*, vol. 22, no. 12, pp. 3762–3767, 2010.

[76] J. Behler, "Atom-centered symmetry functions for constructing high-dimensional neural network potentials," *The Journal of Chemical Physics*, vol. 134, no. 7, p. 074106, 2011.

[77] J. Behler, "Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations," *Phys. Chem. Chem. Phys.*, vol. 13, pp. 17930–17955, 2011.

[78] J. Behler, "First principles neural network potentials for reactive simulations of large molecular and condensed systems," *Angewandte Chemie International Edition*, vol. 56, no. 42, pp. 12828–12840, 2017.

[79] C. M. Handley and P. L. A. Popelier, "Potential energy surfaces fitted by artificial neural networks," *The Journal of Physical Chemistry A*, vol. 114, no. 10, pp. 3371–3383, 2010. PMID: 20131763.

[80] A. R. Oganov and C. W. Glass, "Crystal structure prediction using *ab initio* evolutionary techniques: Principles and applications," *The Journal of Chemical Physics*, vol. 124, no. 24, p. 244704, 2006.

[81] G. Pilania, A. Mannodi-Kanakkithodi, B. P. Uberuaga, R. Ramprasad, J. E. Gubernatis, and T. Lookman, "Machine learning bandgaps of double perovskites," *Scientific Reports*, p. 19375, jan 2016.

[82] M. A. et al., "TensorFlow: Large-scale machine learning on heterogeneous systems." `https://tensorflow.org`, 2015.

[83] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pp. 108–122, 2013.

[84] T. Ahlgren, J. Likonen, J. Slotte, J. Räisänen, M. Rajatora, and J. Keinonen, "Concentration dependent and independent Si diffusion in ion-implanted GaAs," *Phys. Rev. B*, vol. 56, pp. 4597–4603, Aug 1997.

[85] J. J. Murray, M. D. Deal, and D. A. Stevenson, "Influence of background doping and implant damage on the diffusion of implanted silicon in gaas," *Applied Physics Letters*, vol. 56, no. 5, pp. 472–474, 1990.

[86] S. B. Zhang and J. E. Northrup, "Chemical potential dependence of defect formation energies in gaas: Application to ga self-diffusion," *Phys. Rev. Lett.*, vol. 67, pp. 2339–2342, Oct 1991.

[87] J. Dąbrowski and J. E. Northrup, "Microscopic theory of diffusion on the ga sublattice of gaas: Vacancy-assisted diffusion of si and ga," *Phys. Rev. B*, vol. 49, pp. 14286–14289, May 1994.

[88] B. Tuck and M. H. Badawi, "Diffusion of tin in n-type gaas," *Journal of Physics D: Applied Physics*, vol. 11, no. 18, p. 2541, 1978.

[89] D. Shaw, "Alternative mechanisms for the diffusion of sn and zn in gaas," *physica status solidi (a)*, vol. 86, no. 2, pp. 629–635, 1984.

[90] M. B. Panish, "Phase equilibria in the system algaassn and electrical properties of sndoped liquid phase epitaxial alxga1xas," *Journal of Applied Physics*, vol. 44, no. 6, pp. 2667–2675, 1973.

[91] M. E. Greiner and J. F. Gibbons, "Diffusion and electrical properties of silicondoped gallium arsenide," *Journal of Applied Physics*, vol. 57, no. 12, pp. 5181–5187, 1985.

[92] J. Wang, B. Lukose, M. O. Thompson, and P. Clancy, "Ab initio modeling of vacancies, antisites, and Si dopants in ordered InGaAs," *J. Appl. Phys.*, vol. 121, 2017.

[93] A. F. Wright and N. A. Modine, "Migration processes of the as interstitial in gaas," *Journal of Applied Physics*, vol. 120, no. 21, p. 215705, 2016.

[94] J. T. Schick, C. G. Morgan, and P. Papoulias, "First-principles study of as interstitials in gaas: Convergence, relaxation, and formation energy," *Phys. Rev. B*, vol. 66, p. 195302, Nov 2002.

[95] M. Reveil, H.-L. Huang, H.-T. Chen, J. Liu, M. O. Thompson, and P. Clancy, "Ab initio studies of the diffusion of intrinsic defects and silicon dopants in bulk InAs," *Langmuir*, vol. 33, no. 42, pp. 11484–11489, 2017. PMID: 28915733.

[96] S. A. Dayeh, D. Susac, K. L. Kavanagh, E. T. Yu, and D. Wang, "Structural and room-temperature transport properties of zinc blende and wurtzite InAs nanowires," *Advanced Functional Materials*, vol. 19, no. 13, pp. 2102–2108, 2009.

[97] S. Adachi, *Properties of Semiconductor Alloys: Group-IV, III-V and II-VI Semiconductors*. Wiley Online Library, 2009.

[98] H. D. Park, S. M. Prokes, and R. C. Cammarata, "Growth of epitaxial InAs nanowires in a simple closed system," *Applied Physics Letters*, vol. 87, no. 6, p. 063110, 2005.

[99] S. A. Dayeh, E. T. Yu, and D. Wang, "III-V nanowire growth mechanism: V/III ratio and temperature effects," *Nano Letters*, vol. 7, no. 8, pp. 2486–2490, 2007. PMID: 17608541.

[100] K. A. Dick, K. Deppert, T. Mrtensson, B. Mandl, L. Samuelson, and W. Seifert, "Failure of the Vapor-Liquid-Solid mechanism in Au-assisted MOVPE growth of InAs nanowires," *Nano Letters*, vol. 5, no. 4, pp. 761–764, 2005. PMID: 15826124.

[101] M. Y. Bashouti, R. T. Tung, and H. Haick, "Tuning the electrical properties of Si nanowire field-effect transistors by molecular engineering," *Small*, vol. 5, no. 23, pp. 2761–2769, 2009.

[102] S. A. Dayeh, D. P. R. Aplin, X. Zhou, P. K. Yu, E. T. Yu, and D. Wang, "High electron mobility InAs nanowire field-effect transistors," *Small*, vol. 3, no. 2, pp. 326–332, 2007.

[103] S. A. Dayeh, E. T. Yu, and D. Wang, "Transport coefficients of InAs nanowires as a function of diameter," *Small*, vol. 5, no. 1, pp. 77–81, 2009.

[104] F. Ning, L.-M. Tang, Y. Zhang, and K.-Q. Chen, "First-principles study of quantum confinement and surface effects on the electronic properties of InAs nanowires," *Journal of Applied Physics*, vol. 114, no. 22, p. 224304, 2013.

[105] K. Alam and R. N. Sajjad, "Electronic properties and orientation-dependent performance of InAs nanowire transistors," *IEEE Transactions on Electron Devices*, vol. 57, pp. 2880–2885, Nov 2010.

[106] C. L. dos Santos and P. Piquini, "Diameter dependence of mechanical,

electronic, and structural properties of InAs and InP nanowires: A first-principles study," *Phys. Rev. B*, vol. 81, p. 075408, Feb 2010.

[107] S. Wirths, K. Weis, A. Winden, K. Sladek, C. Volk, S. Alagha, T. E. Weirich, M. von der Ahe, H. Hardtdegen, H. Lüth, N. Demarina, D. Grützmacher, and T. Schäpers, "Effect of Si-doping on InAs nanowire transport and morphology.," *Journal of Applied Physics*, vol. 110, no. 5, p. 053709, 2011.

[108] C. Rolland, P. Caroff, C. Coinon, X. Wallart, and R. Leturcq, "Inhomogeneous Si-doping of gold-seeded InAs nanowires grown by molecular beam epitaxy," *Applied Physics Letters*, vol. 102, no. 22, p. 223105, 2013.

[109] H. Ghoneim, P. Mensch, H. Schmid, C. D. Bessire, R. Rhyner, A. Schenk, C. Rettner, S. Karg, K. E. Moselund, H. Riel, and M. T. Björk, "In situ doping of catalyst-free InAs nanowires," *Nanotechnology*, vol. 23, no. 50, p. 505708, 2012.

[110] M. von der Ahe, H. Hardtdegen, K. Sladek, A. Penz, F. Dorn, A. Heiss, T. Weirich, and D. Grützmacher, "Influence of Si-doping on structure in InAs nanowires," *Acta Crystallographica Section A*, vol. 66, p. s300, Sep 2010.

[111] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, "Quantum espresso: a modular and open-source software project for quantum simulations of materials," *J. Phys.: Condens. Matter*, vol. 21, p. 395502, 2009.

[112] K. A. Dick, P. Caroff, J. Bolinsson, M. E. Messing, J. Johansson, K. Deppert, L. R. Wallenberg, and L. Samuelson, "Control of IIIV nanowire crystal structure by growth parameter tuning," *Semiconductor Science and Technology*, vol. 25, no. 2, p. 024009, 2010.

[113] M. Reveil, J. Wang, M. O. Thompson, and P. Clancy, "Preferred diusional pathways of intrinsic defects and silicon dopants in an ordered phase of $In_{0:5}Ga_{0:5}As$: A first-principles study," *Acta Materialia*, vol. 140, pp. 39–45, 2017.

[114] J. E. Northrup and S. B. Zhang, "Dopant and defect energetics : Si in GaAs," *Physical Review B*, vol. 47, pp. 6791–6794, 1993.

[115] C.-W. Lee, B. Lukose, M. O. Thompson, and P. Clancy, "Energetics of neutral Si dopants in InGaAs: An *ab initio* and semiempirical Tersoff model study," *Physical Review B*, vol. 91, p. 0941081, 2015.

[116] S.-G. Lee and K. L. Chang, "Energetics and hydrogen passivation of carbon-related defects in InAs and $In_{0.5}Ga_{0.5}As$," *Physical Review B*, vol. 53, pp. 9784–9790, 1996.

[117] H. A. Jr., A. G. Lind, C. C. Bomberger, Y. Puzyrev, C. Hatem, R. M. Gwilliam, J. M. O. Zide, S. T. Pantelides, M. E. Law, and K. S. Jones, "Implantation and diffusion of silicon marker layers in $In_{0.53}Ga_{0.47}As$," *Journal of Elec. Materi.*, vol. 45, p. 4282, 2016.

[118] H. A. Jr, A. G. Lind, M. E. Law, C. Hatem, , and K. S. Jones, "Continuum modeling of silicon diffusion and activation in $In_{0.53}Ga_{0.47}As$," *ECS Trans.*, vol. 66, p. 57, 2015.

[119] A. E. Mascarenhas, *Spontaneous Ordering in Semiconductor Alloys*. New York: Kluwer Academic/Plenum, 2002.

[120] S.-H. Wei and A. Zunger, "Fingerprints of CuPt ordering in III-V semiconductor alloys: Valence-band splittings, band-gap reduction, and x-ray structure factors," *Phys. Rev. B*, vol. 57, p. 8983, 1998.

[121] T. Pearsall and G. B. Stringfellow, "Compositional modulation and ordering in semiconductors," *MRS Bull.*, vol. 22, p. 16, 1997.

[122] A. Gomyo, T. Suzuki, K. Kobayashi, S. Kawata, I. Hino, and T. Yuasa, "Evidence for the existence of an ordered state in $Ga_{0.5}In_{0.5}P$ grown by metalorganic vapor phase epitaxy and its relation to band-gap energy," *Appl. Phys. Lett.*, vol. 50, p. 673, 1987.

[123] F. E. Fujita, *Physics of New Materials*. Berlin: Springer Berlin Heidelberg, 1998.

[124] O. Ueda, Y. Nakata, T. Nakamura, and T. Fujii, "Tem investigation of cuau-i type ordered structures in MBE-grown InGaAs crystals on (110) InP substrates," *J. Cryst. Growth*, vol. 115, p. 375, 1991.

[125] T. S. Kuan, T. F. Kuech, W. I. Wang, and E. L. Wilkie, "Long-range order in $al_xga_{(1-x)}as$," *Phys. Rev. Lett.*, vol. 54, p. 201, 1985.

[126] H. R. Jen, M. J. Cherng, and G. B. Stringfellow, "Ordered structures in $GaAs_{0.5}Sb_{0.5}$ alloys grown by organometallic vapor phase epitaxy," *Appl. Phys. Lett.*, vol. 48, p. 1603, 1986.

[127] A. Nakano, "A space–time-ensemble parallel nudged elastic band algorithm for molecular kinetics simulations," *Computer Physics Communications*, vol. 178, no. 4, 2008.

[128] H. Jónsson, G. Mills, and K. W. Jacobson, *Nudged Elastic Band Method for Finding Minimum Emergy Paths of Transitions*. World Scientific, 1998.

[129] T. E. Haynes, W. K. Chu, and S. T. Picraux, "Direct measurement of evaporation during rapid thermal processing of capped GaAs," *Appl. Phys. Lett.*, vol. 50, 1987.

[130] A. Alian, G. Brammertz, N. Waldron, C. Merckling, G. Hellings, H. C. Lin, W. E. Wang, M. Meuris, K. D. M. E. Simoen, and M. Heyns, "Silicon and selenium implantation and activation in $In_{0.53}Ga_{0.47}As$ under low thermal budget conditions," *Microelectron. Eng.*, vol. 88, 2011.

[131] M. E. Greiner and J. F. Gibbons, "Diffusion of silicon in gallium arsenide using rapid thermal processing: Experiment and model," *Applied Physics Letters*, vol. 44, no. 8, pp. 750–752, 1984.

[132] M. Reveil and P. Clancy, "Classification of spatially resolved molecular fingerprints for machine learning applications and development of a codebase for their implementation," *Mol. Syst. Des. Eng.*, pp. –, 2018.

[133] L. Ward and C. Wolverton, "Atomistic calculations and materials informatics: A review," *Current Opinion in Solid State and Materials Science*, vol. 21, no. 3, pp. 167 – 176, 2017. Materials Informatics: Insights, Infrastructure, and Methods.

[134] A. Khorshidi and A. A. Peterson, "Amp: A modular approach to machine learning in atomistic simulations," *Computer Physics Communications*, vol. 207, no. Supplement C, pp. 310 – 324, 2016.

[135] H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, "Representa-

tion of intermolecular potential functions by neural networks," *The Journal of Physical Chemistry A*, vol. 102, no. 24, pp. 4596–4605, 1998.

[136] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning," *Phys. Rev. Lett.*, vol. 108, p. 058301, Jan 2012.

[137] K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. von Lilienfeld, K.-R. Mller, and A. Tkatchenko, "Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space," *The Journal of Physical Chemistry Letters*, vol. 6, no. 12, pp. 2326–2331, 2015. PMID: 26113956.

[138] J. Behler and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces," *Phys. Rev. Lett.*, vol. 98, p. 146401, Apr 2007.

[139] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons," *Phys. Rev. Lett.*, vol. 104, p. 136403, Apr 2010.

[140] N. Artrith, A. Urban, and G. Ceder, "Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species," *Phys. Rev. B*, vol. 96, p. 014112, Jul 2017.

[141] L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, "Big data of materials science: Critical role of the descriptor," *Phys. Rev. Lett.*, vol. 114, p. 105503, Mar 2015.

[142] K. Yao, J. E. Herr, and J. Parkhill, "The many-body expansion combined with neural networks," *The Journal of Chemical Physics*, vol. 146, no. 1, p. 014106, 2017.

[143] K. Yao, J. E. Herr, S. N. Brown, and J. Parkhill, "Intrinsic bond energies from a bonds-in-molecules neural network," *The Journal of Physical Chemistry Letters*, vol. 8, no. 12, pp. 2689–2694, 2017. PMID: 28573865.

[144] A. Sadeghi, S. A. Ghasemi, B. Schaefer, S. Mohr, M. A. Lill, and S. Goedecker, "Metrics for measuring distances in configuration spaces," *The Journal of Chemical Physics*, vol. 139, no. 18, p. 184118, 2013.

[145] F. Pietrucci and W. Andreoni, "Graph theory meets ab initio molecular

dynamics: Atomic structures and transformations at the nanoscale," *Phys. Rev. Lett.*, vol. 107, p. 085504, Aug 2011.

[146] L. Zhu, M. Amsler, T. Fuhrer, B. Schaefer, S. Faraji, S. Rostami, S. A. Ghasemi, A. Sadeghi, M. Grauzinyte, C. Wolverton, and S. Goedecker, "A fingerprint based metric for measuring similarities of crystalline structures," *The Journal of Chemical Physics*, vol. 144, no. 3, p. 034203, 2016.

[147] D. Filimonov, V. Poroikov, Y. Borodina, and T. Gloriozova, "Chemical similarity assessment through multilevel neighborhoods of atoms: definition and comparison with the other descriptors," *Journal of Chemical Information and Computer Sciences*, vol. 39, no. 4, pp. 666–670, 1999.

[148] M. Sastry, J. F. Lowrie, S. L. Dixon, and W. Sherman, "Large-scale systematic analysis of 2d fingerprint methods and parameters to improve virtual screening enrichments," *Journal of Chemical Information and Modeling*, vol. 50, no. 5, pp. 771–784, 2010. PMID: 20450209.

[149] P. Willett, "Similarity-based virtual screening using 2d fingerprints," *Drug Discovery Today*, vol. 11, no. 23, pp. 1046 – 1053, 2006.

[150] J. Behler, "Perspective: Machine learning potentials for atomistic simulations," *The Journal of Chemical Physics*, vol. 145, no. 17, p. 170901, 2016.

[151] K. V. J. Jose, N. Artrith, and J. Behler, "Construction of high-dimensional neural network potentials using environment-dependent atom pairs," *The Journal of Chemical Physics*, vol. 136, no. 19, p. 194111, 2012.

[152] A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," *Phys. Rev. B*, vol. 87, p. 184115, May 2013.

[153] G. Ferr, T. Haut, and K. Barros, "Learning molecular energies using localized graph kernels," *The Journal of Chemical Physics*, vol. 146, no. 11, p. 114107, 2017.

[154] D. Rogers and M. Hahn, "Extended-connectivity fingerprints," *Journal of Chemical Information and Modeling*, vol. 50, no. 5, pp. 742–754, 2010.

[155] K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross, "How to represent crystal structures for machine learning: Towards fast prediction of electronic properties," *Phys. Rev. B*, vol. 89, p. 205118, May 2014.

[156] O. A. von Lilienfeld, R. Ramakrishnan, M. Rupp, and A. Knoll, "Fourier series of atomic radial distribution functions: A molecular fingerprint for machine learning models of quantum chemical properties," *International Journal of Quantum Chemistry*, vol. 115, no. 16, pp. 1084–1093, 2015.

[157] C. R. Collins, G. J. Gordon, O. A. von Lilienfeld, and D. J. Yaron, "Constant size molecular descriptors for use with machine learning," 2016.

[158] V. Botu and R. Ramprasad, "Learning scheme to predict atomic forces and accelerate materials simulations," *Phys. Rev. B*, vol. 92, p. 094306, Sep 2015.

[159] T. D. Huan, R. Batra, J. Chapman, S. Krishnan, L. Chen, and R. Ramprasad, "A universal strategy for the creation of machine learning-based atomistic force fields," *npj Computational Materials*, vol. 3, pp. 89 – 109, 217.

[160] M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzsenyi, and P. Marquetand, "Wacsf - weighted atom-centered symmetry functions as descriptors in machine learning potentials," *ArXiv e-prints*, dec 2017.

[161] B. Huang and O. A. von Lilienfeld, "Communication: Understanding molecular representations in machine learning: The role of uniqueness and target similarity," *The Journal of Chemical Physics*, vol. 145, no. 16, p. 161102, 2016.

[162] H. Huo and M. Rupp, "Unified representation for machine learning of molecules and crystals," *ArXiv e-prints*, apr 2017.

[163] S. Kearnes, K. McCloskey, M. Berndl, V. Pande, and P. Riley, "Molecular graph convolutions: moving beyond fingerprints," *Journal of Computer-Aided Molecular Design*, vol. 30, pp. 595–608, Aug 2016.

[164] O. Isayev, C. Oses, C. Toher, E. Gossett, S. Curtarolo, and A. Tropsha, "Universal fragment descriptors for predicting properties of inorganic crystals," *Nature Communications*, vol. 8, p. 15679, Jun 2017.

[165] S. M. Kandathil, T. L. Fletcher, Y. Yuan, J. Knowles, and P. L. A. Popelier, "Accuracy and tractability of a kriging model of intramolecular polarizable multipolar electrostatics and its application to histidine," *Journal of Computational Chemistry*, vol. 34, no. 21, pp. 1850–1861, 2013.

[166] T. D. Huan, A. Mannodi-Kanakkithodi, and R. Ramprasad, "Acceler-

ated materials property predictions and design using motif-based fingerprints," *Phys. Rev. B*, vol. 92, p. 014106, Jul 2015.

[167] J. P. Janet and H. J. Kulik, "Resolving transition metal chemical space: Feature selection for machine learning and structureproperty relationships," *The Journal of Physical Chemistry A*, vol. 121, no. 46, pp. 8939–8954, 2017. PMID: 29095620.

[168] J. P. Janet and H. J. Kulik, "Predicting electronic structure properties of transition metal complexes with neural networks," *Chem. Sci.*, vol. 8, pp. 5137–5152, 2017.

[169] B. Jiang and H. Guo, "Permutation invariant polynomial neural network approach to fitting potential energy surfaces," *The Journal of Chemical Physics*, vol. 139, no. 5, p. 054112, 2013.

[170] J. Li, B. Jiang, and H. Guo, "Permutation invariant polynomial neural network approach to fitting potential energy surfaces. ii. four-atom systems," *The Journal of Chemical Physics*, vol. 139, no. 20, p. 204103, 2013.

[171] R. Kondor, "A complete set of rotationally and translationally invariant features for images," *CoRR*, vol. abs/cs/0701127, 2007.

[172] A. Thompson, L. Swiler, C. Trott, S. Foiles, and G. Tucker, "Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials," *Journal of Computational Physics*, vol. 285, no. Supplement C, pp. 316 – 330, 2015.

[173] M. Novotni and R. Klein, "Shape retrieval using 3d zernike descriptors," *Computer-Aided Design*, vol. 36, no. 11, pp. 1047 – 1062, 2004. Solid Modeling Theory and Applications.

[174] N. Artrith and A. Urban, "An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for tio2," *Computational Materials Science*, vol. 114, no. Supplement C, pp. 135 – 150, 2016.

[175] K. Yao, J. E. Herr, D. W. Toth, R. Mcintyre, and J. Parkhill, "The tensormol-0.1 model chemistry: a neural network augmented with long-range physics," *ArXiv e-prints*, nov 2017.

[176] S. Fritsch and F. Guenther, *neuralnet: Training of Neural Networks*, 2016. R package version 1.33.

[177] R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.